



Cloning Scheme for Generating pCMV-NS35

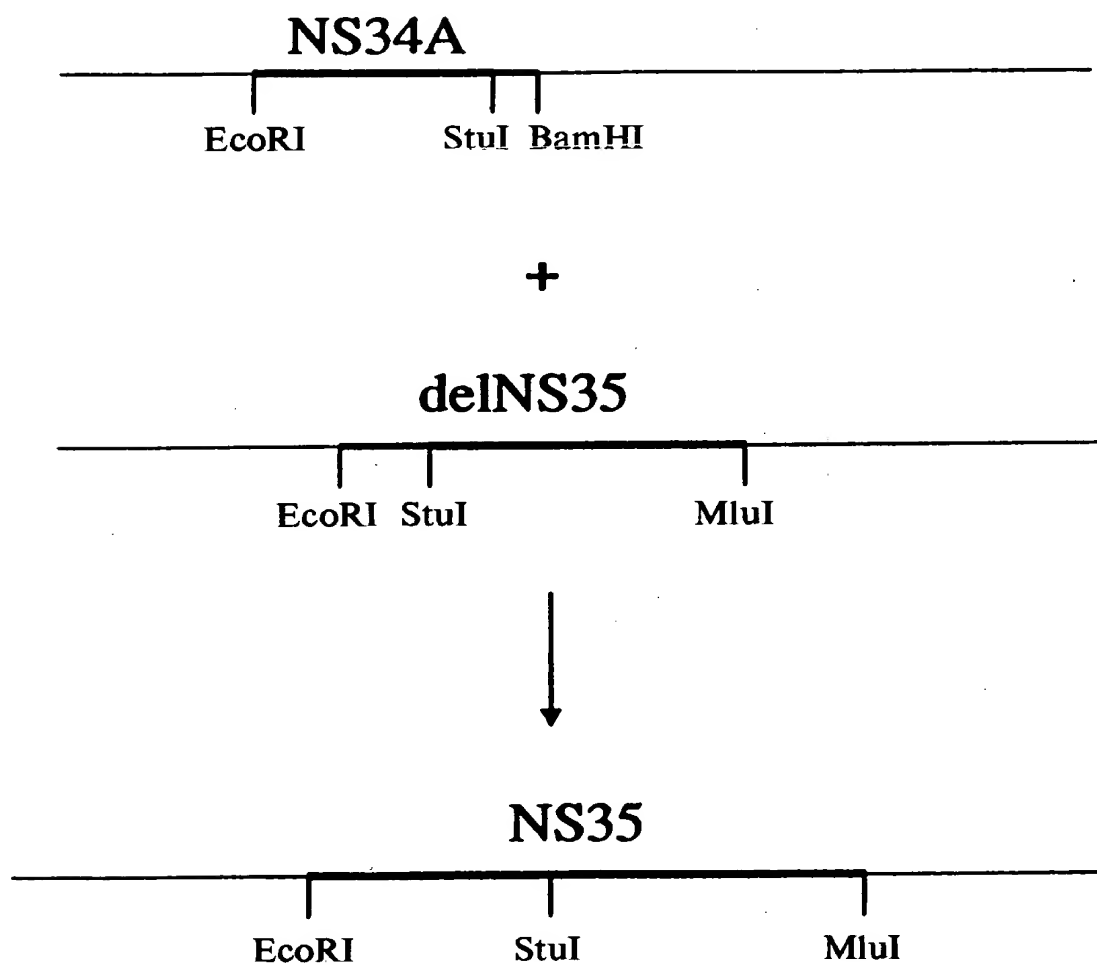


FIG. 1

09/721479

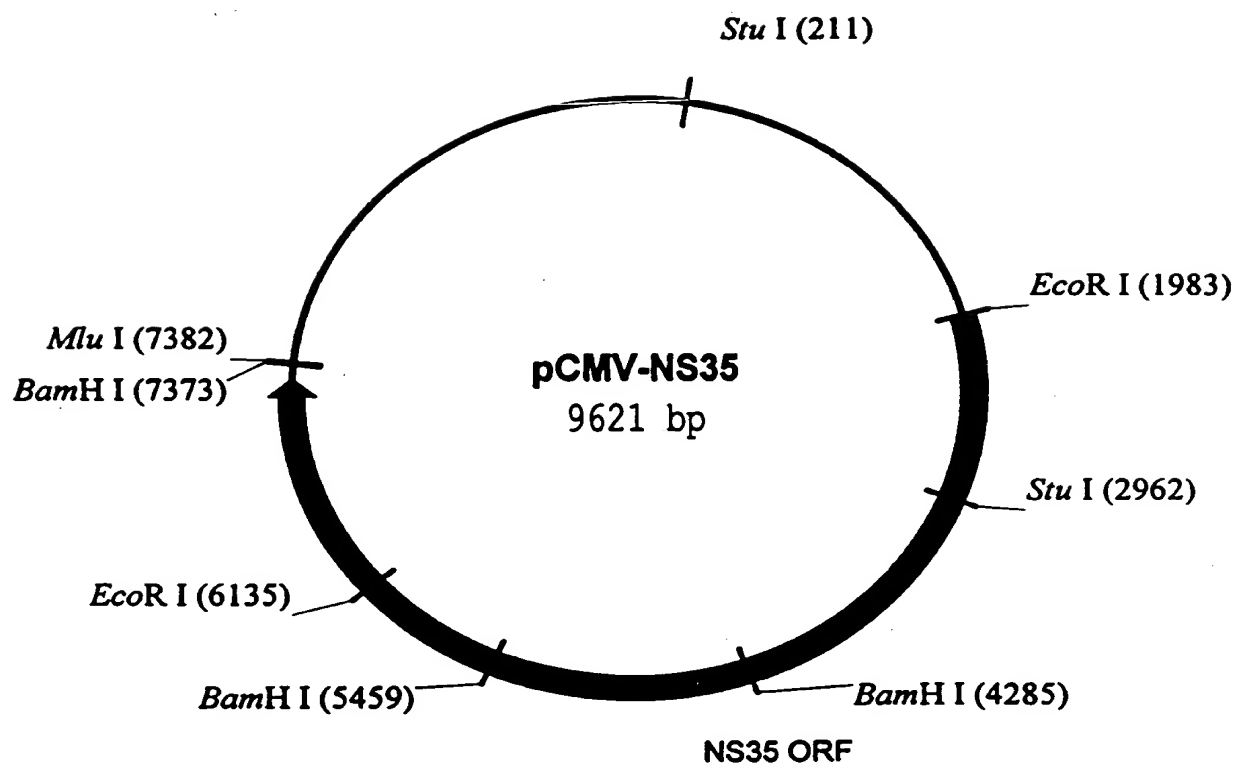


FIG. 2

pCMV-NS35

1	TCGCGCGGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCGCTA
81	GCGGGGAGCA GACAAGCCCG TCAGGGGGCG TCAGGGGGTG TTGGGGGGTG TCGGGGGCTGG CTTAACTATG CGGCATCAGA CGGCCCTCGT CTGTTGGGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACIT CGAAAAACGT TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGGCGGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGGGGAG AATGGCGGGA TTATCGAGTC TCGGGCTCGG CCGGAGCGCG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGGCTC TTACCCGCTC
321	ACTGGGGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTTCA GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA
481	AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGGCCCAT TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGGGGTTAT CCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA
641	AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTTACTGCC ATTTACCGGG

StuI

FIG. 3-1



pCMV-NS35

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACGGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCAGGGG ATTTCCAAGT CTCCACCCCA
GTACCACTAC GCCAAAACCG TCATGTGGT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGT

881 TTGACGCTAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAAACCCGC CCCGTTGACG
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
GTTTACCGCG CATCCGCACA TGCCACGCTC CAGATATATT CGTCTGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC
GGTAGGTGG ACAAACCTGG AGGTATCTTC TGTGGCCCTG CTTAGGTCGG AGGGCGGGC CCTTGCCACG TAACCTTGCG

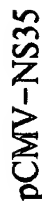
1121 GGATTCCCG TGCCAAGAGT GACGTAAGTA CCGCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGCG ACGGTTCTCA CTGCATTCTAT GCGGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
GACAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATT TTAGCCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGG GTCCATTAT
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGCTCTACCC CAGGTAAATA

FIG. 3-2



+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGCT TTGGTGCTTA CATGTCCAAG
TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTTC

၆၆

pCMV-NS35

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
 2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGGTACT CCACCTACGG
 CGAGTACCCT AGCTAGGATT GTAGTCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
 2161 CAAGTTCTTT GCCGACGGG GGTGCTCGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
 GTTCAAGGAA CGGCTGCCG CCACGAGCCC CCCGGAATA CTGTATTATT AACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
 2241 CATCCATCTT GGGCATTGG ACTGTCTTG ACCAAGCAGA GACTGCGGG GCGAGACTGG TTGTGCTCG CACCGCCACC
 GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
 2321 CCTCGGGCT CGGTCACTGT GCCCATCCC AACATCGAG AGTTGTCT CTCCACCACC GGAGAGATCC CTTTTACGG
 GGAGGCCGA GGCAGTGACA CGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
 2401 CAAGGCTATC CCCCTCGAAG TAATCAAGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAACCTG
 GTCCGATAG GGGAGCTTC ATTAGTTCCC CCCCTCTGA GAGTAGAGA CAGTAAGTTT CTTCTTCAGG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
 2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
 GGGTTTGA CCAGCTAAC CCGTAGTTAC GGCACCGGAT GATGGGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCC

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
 2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG
 CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

FIG. 3-4



pCMV-NS35

```

+2  V T Q T V D F S L D P T F T I E T I T L P Q D A V S
2641 TGTACCCAG ACAGTCGATT TCAGCCCTTGA CCCTACCTTC ACCATTGAGA CAATCAGCT CCCCCAAGAT GCTGTCTCCC
    ACAGTGGGTC TGTACGCTAA AGTCGGAAG AGTCGGAAGT GGGATGGAAG TGGTAACCTCT GTTAGTGGGA GGGGGTTCTA GCACAGAGGG

+2  R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
    CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCCGTG GCCCCCTCGC GGGGAGGCGC

+2  M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
2801 ATGTTGACT CGTCCGTCCT CTGTGAGTGC TATGACGAG GCTGTGCTTG GTATGAGTCT ACGCCGCGCG AGACTACAGT
    TACAAGCTGA GCAGGCAGGA GAACTCAGG ATACTGCGTC GCACACGAAC CATACTCGAG TCGGGGCGCG TCTGATGTCA

+2  R L R A Y M N T P G L P V C Q D H L E F W E G V F T
    StuI

2881 TAGGTACGA GGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
    ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTTCCG CAGAAATGTC

+2  G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
    StuI
    ----
2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAC
    CGGAGTGAGT ATATCTACGG GTGAAAGATA GGTCTGTTT CGTCTCACC CGTCTGGAAG GAATGGACCA TCGCATGGTT

+2  A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCTT CCCCCCTCGT GGGACAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCAC
    CCGTGGCACA CGCGATCCCG AGTTCGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG

```

FIG. 3-5



pCMV-NS35

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCCCT GACGCACCCA GTCACCAAAAT
GGAGGTACCC GGTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTAGTGGA CTGGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201 ACATCATGAC ATGCATGTC GCGACCTGG AGGTGCTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG
TGTAAGTCTG TAGGTACAGC CGGTGGACC TCCAGCAGTG CTCGTGGACC CAGGAGCAAC CGCCGCAGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281 GCCGGTATT GCCTGTCAAC AGGTGGCTG GTCATAGTGG GCAGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
CGCGCATAA CGACAGTGT TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTGGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTGG TGAATGGCAT GTAGTCTGTT CCCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGGTCC CGTCAGGCAG AGTTATCGC CCCTGCTGTC
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGGCGCAG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521 CAGACCAACT GGCACCAACT CGAGACCTTC TGGGGAAGC ATATGTGGA CTTCATCAGT GGGATACAAT ACTTGGCGGG
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCGGTTG TATACACCTT GAAGTACTCA CCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 CTGTCAACG CTGCCTGGTA ACCCGCCAT TGCTTCATG ATGGCTTTTA CAGTGTGTGT CACGAGCCCA CTAACCACTA
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGAGGACA GTGGTCCGGT GATTGGTGAT

FIG. 3-6



pCMV-NS35

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 GCCAAACCCT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCGGCCCGG GTGCCGCTAC TGCCTTTGTG
CGGTTTGGGA GGAGAAGTIG TATAACCCCG CCACCCACCG ACGGCTCGAG CCGCGGGGGC CACGGCGATG ACGGAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGCTGGCT TAGCTGGCG CGCCATCGG AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCG
CCGCGACCGA ATCGACCGG GCGTAGCCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GCGCTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
CCGCGACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGTGCCT CCTGGACCAG TTAGATGAGC

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
3921 CCGCCATCCT CTCGCCCGGA GCCCTCGTAG TCGCGTGGT CTGTGCAGCA ATACTGCGC GGCACGTTGG CCGGGGGGAG
GGCGTAGGA GAGCGGGCCT CCGGAGCATC AGCCGACCA GACACGTCGT TATGACGCG CCGTGAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTGCCT CCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
CCCCGTCAGC TCACCTACTT GGCGGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAG GGTGCGCTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGCGATGCA GCTGCCCGG TCACTGCCAT ACTCAGCAGC CTCAGTGTA CCCAGCTCCT GAGCGGACTG CACCACTGGA
CTCGTACGT CGACGGGGC AGTGACGTA TGATCGTCG GAGTGACATT GGTGCGAGGA CTCCGCTGAC GTGCTCACT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
4161 TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC
ATTGAGCCT CACATGGTGA GGTACGAGG CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

FIG. 3-7



pCMV-NS35

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
BamHI

4241 TTAAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG
AAATTCTGGA CCGATTTTTCG ATTGAGTAC GGTGTGACG GACCTAGGG GAAACACAGG ACGTGGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T
4321 GGTCTGGGA GGGACGGCA TCATGCACAC TCGTGGCAGC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA
CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C
4401 TGAGGATCGT CGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGACCTTC CCCATTAAATG CCTACACCAC GGGCCCTGT
ACTCTAGCA GCCAGGATCC TGGACGTCTT TGTACACCTC ACCCTGGAAG GGGTAAATTAC GGATGTGGTG CCGGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
4481 ACCCCCTTC CTGCGCGGA CTACACGTC GCGTATGGA GGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGG
TGGGGGAAG GACGGGGCTT GATGTGCAAG CCGGATACCT CCCACAGAGG TCTCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATGCCCCGAA TTTTTCACAG
CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTAGAA TTACGGGCA CCGTCCAGG TAGCGGGCTT AAAAAGTGC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G
4641 AATTGGACGG GGTGGCCTA CATAGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA
TTAACCTGCC CCACGGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCTT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCCACGAAT ACCCGGTAGG GTCCGAATTA CCTTGGGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GCTTGGCTT GCACCGGCAC AACTGCAGGT ACGAGTGACT

FIG. 3-8



pCMV-NS35

+2 P S H I T A E A A G R R L A R G S P P S V A S S A
4801 TCCCTCCCAT ATAACAGCAG AGCGGCGG GCGAAGTTG GCGAGGGAT CACCCCTC TGTGGCCAGC TCCTCGGCTA
AGGAGGGTA TATTGTCGTC TCCGCGGCGC CGCTTCCAAC CGTCCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC
CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGTAATG AGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGT GTGATTCTGG ACTCCTTCGA
GAGGATACCT CCGTCTCTTA CCCGCCGTTG TAGTGTCCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

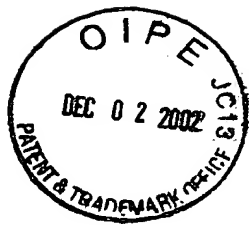
+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCGCTGTG GCGAGGAGG ACGAGCGGA GATCTCCGA CCGCAGAAA TCCTCGGAA GTCTCGGAGA TTCGCCAGG
AGGCGAACAC CGCTCTCTCC TGCTCGCCCT CTAGAGGAT GGGGCTCTT AGGACGCTT CAGAGCCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCCTGCCCCG TTGGGCGCG CCGGACTATA ACCCCCGCT ACTGAGACG TGGAAGGAG CCGACTACGA ACCACCTGTG
GGGAGGGCA AACCCGCGG GGCCTGATAT TGGGGGGCGA TCACTCTGC ACCTTTTTCG GGTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E
5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGC TCAGGAGAG CGGACGGTGG TCCTCACTGA
CAGGTACCGA CGGGCGAAG TGGAGTTTC AGGGAGGAC ACGGAGCGG AGCTTCTTC GCCTGCCACC AGGAGTGA

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D
5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCTCAAC TTCGGGCAAT ACGGGCGACA
TAGTTGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

FIG. 3-9



pCMV-NS35

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTTT CTGGCTGCCC CCGGACTCC GACGCTGAGT CCTATTCTC CATGCCCCC
TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGGCACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V
BamHI

5441 CTGGAGGGG AGCCTGGGA TCCGGATCTT AGCGACGGGT CATGGTCAAC GGTCACTAGT GAGGCCAACG CGGAGGATGT
GACCTCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCC GTACCAGTTG CCAGTCATCA CTCGGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGGCACTC GTCACCCCGT GCGCGCGGA AGAACAGAAA CTGCCCCATCA
GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCTGAG CAGTGGGCA GCGGGCGCCT TCTTGCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCACTAAG CAATCGTTG CTACGTACC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAAG
TACGTGATTC GTTGAACAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTGCTGTA TGGTCTGCA TGAGTTCTC CAATTTCGTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGCAGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTGTTATG
TCACCTCCGA TTGAACGATA GGCATCTCCT TCGAACGTG GACTGGGGG GTGTGAGTGG GTTAGGTTT AAACCAATAC

FIG. 3-10



PCMV-NS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGCAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAA CACATCAAC TCCGTGTGA AAGACCTTCT GGAAGACAAT
CCCGTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATG GGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC
CATTCTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGAGGCA AGTCGGACTC TTCCCCCAG CATTGGGTGG

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCGTCTCATC GTGTTCCTCG ATCTGGCGGT GCGCGTGTGC GAAAGATGG CTTTGTACGA CGTGGTTACA AAGTCCCT
AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTCTTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
ACCGGCACTA CCTTCGAGG ATGCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAGG AGCAGGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAACCC CAATGGGTTT CTCGTATGAT ACCCGTGTCT TTGACTCCAC AGTCACCTGAG AGGCACATCC GTACGGAGGA
TTCTTTTGGG GTTACCCCAA GACATACTA TGGCGGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
CCGTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGGC CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCGC GCGAGGGCGG TACTGACAAC TAGCTGTGGT
GGGGAAGATG GTTAAGTTCC CCCCTCTTGA CGCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTG ATCGACACCA

FIG. 3-11



pCMV-NS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G
6401 AACACCCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCG CAGG GCTCCAGGAC TGCACCATGC TCGTGTCTGG
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTG ACGTGGTAGG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCAGG GAGGCTATGA
GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCAGGTC CTCCTGGCGG GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
6561 CCAGGTACTC CGCCCCCCT GGGACCCCG CACAACCGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641 TCAGTGGCC ACGACGGCGG TGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGGA GAGCTGCGTG
AGTCAGCGG TGCTGCGCGG ACCTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTCGCCCC CACACTGTGG GCGAGGATGA
CCTCTGCGT TCTGTGTGAG GTCACTTAAG GACCGATCCG TTGTATTAGT ACAACGCGG GTGTGACACC CGTCTCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
6801 TACTGATGAC CCATTCTTTT AGCGTCCTTA TAGCCAGGA CCAGCTTGAA CAGGCCCTCG ATTGGGAGAT CTACGGGGCC
ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGTCCCT GTCCGAACCT GTCCGGGAGC TAACGCTCTA GATGCCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881 TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCAATC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
ACGATGAGGT ATCTTGGTGA CCTAGATGA GGTAGTAGT TTTCTGAGT ACCGGAGTCC CGTAAAGTG AGGTGTCAAT

FIG. 3-12



pCMV-NS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
6961 CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG
GAGAGTCCA CTTTAGTTAT CCCACCGGGG TACGGAGTCT TTTGAACCCC ATGGCGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
GGGCTCGCA GCGCGGATCC GAAGACCGGT CTCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAATT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
7121 AGAACAAAGC TCAAACCTCAC TCCAATAGCG GCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
TCTTGTTTCG AGTTGACTG AGTTATCGC GCGCGACCGG TCGACCTGAA CAGGCCGACC AAGTCCCGAC CGATGTCGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCGCG TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG
CCCTCTGTAA ATAGTGTCCG ACAGAGTACG GGCCCGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCGCCATC

+2 G I Y L L P N R
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAAATCTAG AAAGCGCGCG
CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTATAGATC TTTCGGCGCG

BAMHI MluI

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT
GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCC GTAGACAACA

7441 TTGCCCCCTCC CCGGTGCCTT CCTTGACCTT GGAAGTGCC ACTCCCACTG TCCTTTCCTA ATAAATGAG GAAATTGCAT
AACGGGGAGG GGGCACGGAA GGAAC TGGA GGAAGGTGAC AGGAAAGGAT TATTTTACTC CTTTAAACGTA

FIG. 3-13



pCMV-NS35

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTTC TGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGCTTCC CCCTCCTAAC CCTTCTGTTA

7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTGCGTCACT GACTGGCTGC GCTGGTCTGT TCGGCTGCGG CGAGCGGTAT
TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCGACGCC GCTGCCATA

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
GTCCAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAATAA
GTTTCCGGT CCTTGGCATT TTCCGGGCGC AAGGACGCA AAAAGGTATC CGAGGCGGG GGAAGTGTCT TAGTGTCTTT

7841 TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC
AGCTCGGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGAT ATTTCTATGG TCCGCAAGG GGGACCTTCC AGGAGCACG

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACTGTC CGCTTTCTC CCTTCGGAA GCGTGGCGCT TTCTCAATGC
CGAGAGGACA AGCTGGGAC GCGCAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCCGT CCAAGCTGG CTGTGTGCAC GAACCCCGCG TTCAGCCCGA
AGTGGACAT CCATAGAGTC AAGCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT

8081 CCGCTCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
GGCGACGGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCAATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTAG CTACACTAGA
CATTGTCTTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

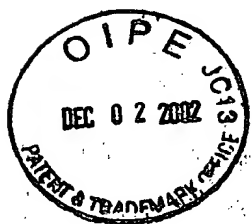
FIG. 3-14



pCMV-NS35

8241	AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA TCCTGTGATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCCTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT
8321	AACCAACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTTCGC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA
8401	TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCCTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC
8481	ATCTTCACCT AGATCCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG TAGAAGTGGA TCTAGGAAAA TTTTAATTTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC
8561	TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGCAGC
8641	TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACGGG GTACGACGCT TACTATGGCG CTCTGGGTGC GAGTGGCCGA
8721	CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGGCAGAAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GGTCTAAATA GTCGTTATTT GGTCGGTCCG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAATAGGC GGAGGTAGGT
8801	GTCTATTAT TGTTCGCGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTGTCG ATTGCTACAG CAGATAATTA ACAACGGCCC TTCGATCTCA TTCAATCAAGC GGTCATTTAT CAAACGCGTT GCAACAACGG TAACGATGTC
8881	GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGCGGAGT TACATGATCC CGTAGCACCA CAGTGGGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGTTGCTA GTTCCGCTCA ATGTACTAGG

FIG. 3-15



pCMV-NS35

8961	CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA
9041	CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTGATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT
9121	CCAAGTCATT CTGAGAAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCTATTATG GCGCGGTGTA
9201	AGCAGAACTT TAAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGGAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTCGAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG
9281	CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCGC CGTTTCTGGG TGAGCAAAAA GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT
9361	CAGGAAGGCA AAATGCCGCA AAAAAAGGAA TAAGGGCGAC ACGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT GTCCCTCCGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA
9441	TATTGAAGCA TTTATCAGG TTATTGTCTC ATGACCGGAT ACATATTGA ATGTATTAG AAAAATAAAC AAATAGGGGT ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAAATC TTTTATTG TTTATCCCCA
9521	TCCGGGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC AGGCGGTGT AAAGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG
9601	GTATCAGGAG GCCCTTTCTG C CATAGTGCTC CGGAAAGCA G

FIG. 3-16

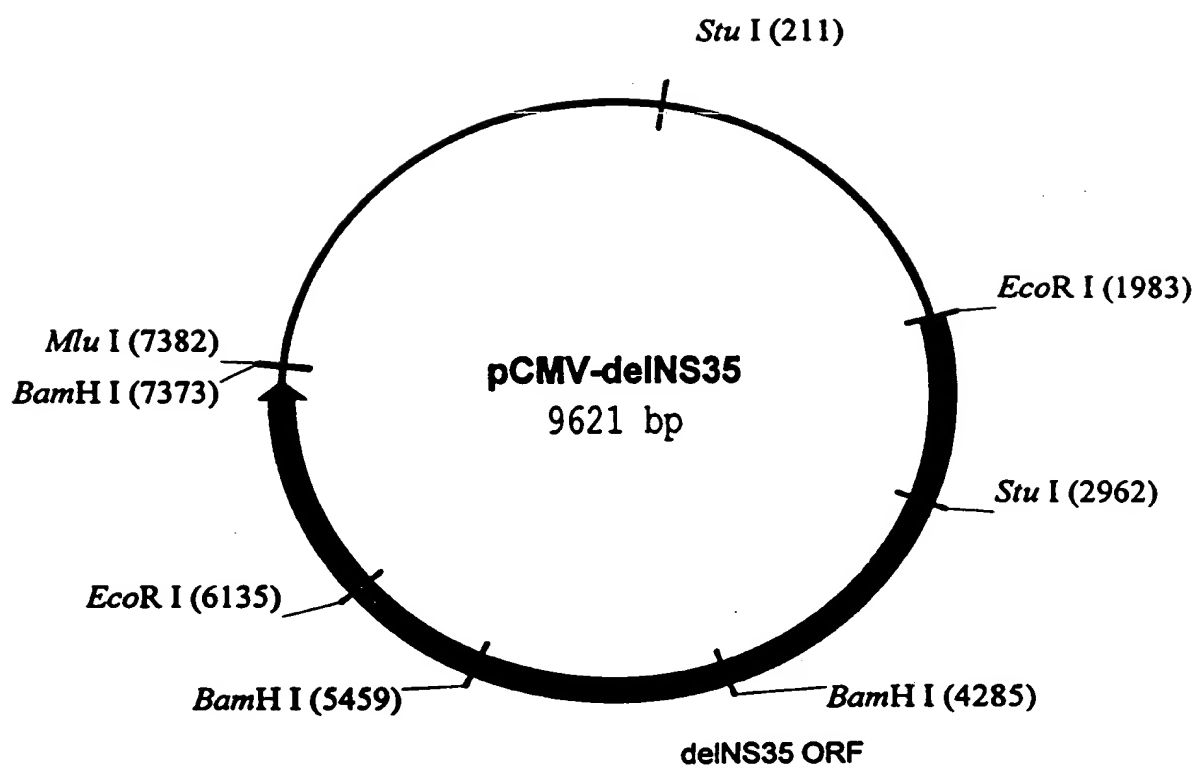
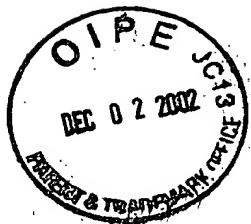


FIG. 4



pCMV-deINS35

1	TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAAAGCGGAT AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGCG CTCTGCCAGT CTCGAACAGA CATTGGCCTA
81	GCCGGGAGCA GACAAGCCCG TCAGGGGCGG TCAGCGGGTG TTGGCGGGTG TCGGGGGTGG CTTAACATATG CGGCATCAGA CGGCCCTCGT CTGTTGGGC AGTCCGCGC AGTCGCCAC AACGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTCGA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGTTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGCCGAGGC GGCCTCGGC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG ATTGGGCGGA TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGAGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCCT
321	ACTGGGCGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT TGACCCGCGC CTCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTTTCA GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA
481	AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT TCGGGTATAT ACCTCAAGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGG GGGTTGCTGG GCGCGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGTATCA TTGGGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

StuI

FIG. 5-1



pCMV-delNS35

641	AAACTGCCCA	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCGGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGGCGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG
721	GCCTGGCAAT	ATGCCCAGTA	CATGACCTTA	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	CGGACCGTAA	TACGGGTTCAT	GTACTGGAAT	GGCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA
	GTACCACTAC	GCCAAAACCG	TCAATGTGTT	ACCGGCACCT	ATCGCCAAAC	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT
881	TTGACGTCAA	TGGGAGTTTG	TTTTTGGCACC	AAAATCAACG	GGACTTTTCCA	AAATGTCTGA	ATAACCCCGC	CCCCTTGACG
	AACTGCAGTT	ACCCTCAAAC	AAAACCGTGG	TTTTTAGTTGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG	GGGCAACTGC
961	CAAAATGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG
	CTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC
1041	CCATCCACGC	TGTTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGGCG	GGAACGGTGC	ATTGGAACGC
	GGTAGGTGCG	ACAAAACCTGG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTGCG	AGCGCCCGGC	CCTTGCCACG	TAACTTTGCG
1121	GGATTCCCGG	TGCCAAGAGT	GACGTAAGTA	CGGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	CCTAAGGGGC	ACGGTTCTCA	CTGCATTTCAT	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CCTACCGATAT
1201	CTGTTTTTTGG	CTTGGGGCCT	ATACACCCGC	GCTCCTTATG	CTATAGGTGA	TGGTATAGCT	TAGCCTATAG	GTGTGGGTTA
	GACAAAACCC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT	ACCATATCGA	ATCGGATATC	CACACCCCAAT
1281	TTGACCATTAA	TTGACCACTC	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG	CCACAACCTAT
	AACTGGTAAT	AACTGGTGAG	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC	GCTGTTTGATA

FIG. 5-2



pCMV-delINS35

1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTACAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT	
	GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA	
1441	TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT	
	ATAAATGTTT AAGTGATAT GTTGTTCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCTT AGAGGCTGTA	
1521	CTCGGGTACG TGTTCGGGAC ATGGGCTCTT CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA	
	GAGCCCATGC ACAAGGCCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT	
1601	GCGGCTCATG GTCGCTCGGC AGTCTCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC	
	CGCCGAGTAC CAGCGAGCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACCG GTGGTGGTGG	
1681	AGTGTCCCGC ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT	
	TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA	
1761	GGAAGACTTA AGGCAGCGC AGAAGAAGAT GCAGGCAAGT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT	
	CCTTCTGAAT TCCGTCCGGC TCTTCTTCTA CGTCCGTCCA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA	
1841	TGCGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGGCGG CGCCACCAGA CATAATAGCT	
	ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGCTCT GTATTATCGA	
+2		M A A
		EcoRI

1921	GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT	
	CTGCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA	

FIG. 5-3



pCMV-delINS35

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTGCTGCA ACACTGGCT TTGTGCTTA CATGTCCAAG
TAGGTCGAGT CCGGATATTC CAGGATCATG AGTTGGGAG ACAACGACGT TGTGACCCGA AACCAACGAAT GTACAGGTTT

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGCTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGCTG ACCGTCGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCGACGGG GGTGCTCGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCG CCACGAGCCC CCCGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCCATCTT GGGCATTTGG ACTGTCTTG ACCAAGCAGA GACTCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTGCTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCCGGGCT CCGTCACTGT CCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCCA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCAATCAA GAAGAAGTGC GACGAACCTCG
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTCTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGGGCCA GAACTGCACA GGCAGTAGGG CTGTCGCGCG

FIG. 5-4



pCMV-delNS35

```
+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
2561 GATGTTGCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGCACTCG GTGATAGACT GCAATACGTC
CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
2641 TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCTACCTTC ACCATTGAGA CAATCAGCT CCCCCAAGAT GCTGTCTCCC
ACAGTGGGTC TGTACGCTAA AGTCGGAAC TGGATGGAAG TGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTGGGTCC GTAGATGCT AAACACCGTG GCCCCCTCGC GGGGAGGCCG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
2801 ATGTTGCACT CGTCCGTCTT CTGTGAGTGC TATGAGGCAG GCTCTGCTTG GTATGAGCTC ACGCCGCGCG AGACTACAGT
TACAAGCTGA GCAGGCAGGA GAACTCAGG ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
Stul

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTCCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
Stul
----
2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGTT
```

FIG. 5-5



pCMV-delINS35

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGGCC TCAAGCCCAC
CGGTGGCACA CGGATCCCG AGTTGGGGG GGGGTAGCA CCTGGTCTA CACCTTACA AACTAAGCGG AGTTGGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGGGCT GTTCAGATG AAATCACCTT GAGGCACCCA GTCACCAAAT
GGAGGTACCC GGTGTGGG ACGATATGTC TGACCCGGGA CAAGTCTTAC TTATGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201 ACATCATGAC ATGCATGTG GCGACCTGG AGTCTGTAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG
TGTAAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281 GCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCTG CTGTCCGGG AAGCCGGCAA TCATACCTGA
CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCAC CGTCCAGCA GAACAGGCC TTGCGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361 CAGGGAAGTC CTCTACCGAG AGTTGGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
GTCCCTTCAG GAGATGGGTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCTCC CGTCAGGCAG AGTTATCGC CCCTGCTGTC
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521 CAGACCAACT GGCAAAACCT CGAGACCTTC TGGGCGAAGC ATATGTGGA CTTCATCAGT GGGATACAAT ACTTGGCGGG
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

FIG. 5-6

pCMV-delINS35

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 CTTGTCAACG CTGCTGGTA ACCCGGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACAGAACTAC TACCGAAAT GTGACGACA GTGGTCGGT GATTGGTGAT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 GCCAAACCCT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCGCCCCCG GTCCGCTAC TGCCTTTGTG
CGGTTTGGGA GGAGAACTTG TATAACCCCC CCACCCACCG ACGGTCTGAG CCGCGGGGGC CACGGCGATG ACGGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGGTGGCT TAGCTGGCG CGCCATCGC AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCTTGCAG GGTATGGCGC
CCGGACCGA ATCGACCGG GCGGTAGCG TCACAACTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GGGCTGGCG GGAGCTCTTG TGGCAITCAA GATCATGAGC GGTGAGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
CCCGACCGC CCTCGAGAAC ACCGTAAGT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
3921 CCGCCATCCT CTCGCCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGGGAG
GGCGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTGCCT CCCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
CCCCGTACG TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GTACAAAGG GGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGGATGCA GGTGCCGCG TCACCTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGTCTCT GAGGCGACTG CACCAGTGGA
CTCGTACGT CGACGGGCG AGTGACGGTA TGAGTCGTCG GAGTGACATT GGTGCGAGGA CTCGCTGAC GTGGTCACCT

FIG. 5-7



pCMV-deINS35

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
4161 TAAGCTCGGA GTGTACCACT CCATGCTCG GTTCTGGCT AAGGACATC TGGACTGGA TATGCGAGT GTTGAGCGAC
ATTGAGCCT CACATGGTGA GGTACGAGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACCTGGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G

BamHI

4241 TTAAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG
AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTCTGAGC GACCCTAGG GAAACACAGG ACGTGGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T
4321 GGTCTGGCGA GGGACGGCA TCATGCCAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA
CCAGACCGCT CCCCTGCCGT ACTACGTGTG ACGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

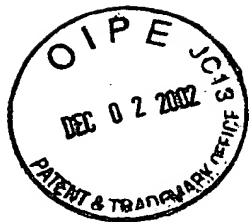
+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C
4401 TCAGGATCGT CGTCTCTAGG ACCTGCAGGA ACATGTGGAG TGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCCTGT
ACTCCTAGCA GCCAGGATCC TGGACGTCT TGTACACCTC ACCCTGGAAG GGTAATTAC GGATGTGGTG CCGGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
4481 ACCCCCCTTC CTGCGCCGAA CTACACGTTG GCGCTATGCA GGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG
TGGGGGGAAG GACGGGGCTT GATGTGCAAG CGGATACCT CCCACAGAGC TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATCGCCCGAA TTTTTCACAG
CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTACGGGCA CGTCCAGG TAGCGGGCTT AAAAAGTGC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G
4641 AATTGGACGG GGTGGCCCTA CATAGGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA
TTAACCTGCC CCACGGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

FIG. 5-8



pCMV-deINS35

```
+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCCACGAAT ACCGGTAGG GTCGCAATTA CTTGGCAGC CCGAACCGGA GTGGCCGTG TTGACGTCCA TGCTCACTGA
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A
4801 TCCCTCCCAT ATAACAGCAG AGCGGGCCGG CGCAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA
AGGAGGGTA TATTGTCGTC TCCGCCGGGC CGCTTCCAAC CGCTCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC
CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGGA GCGAGGAGAT GGGCGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA
GAGGATACCT CCGTCTCTA CCGCCGTTG TAGTGTTCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCGCTTGTG GCGGAGGAGG ACCAGCGGGA GATCTCCGTA CCGCAGAAA TCCTGGGAA GTCTCGGAGA TTCGCCCAGG
AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCCT CAGAGCCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCCTGCCCCG TTGGCGCGG CCGGACTATA ACCCCCGCT AGTGGAGACG TGGAAAAGC CCGACTACGA ACCACCTGTG
GGGACGGGCA AACCCGGGCG GGCTGATAT TGGGGGGGA TCACCTCTGC ACCTTTTTC GGCTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E
5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCTCCTG TGCCTCGGC TCAGGAAGAAG CGGACGGTGG TCCTCACTGA
CAGGTACCGA CCGGCGAAGG TGGAGTTTC AGGGAGGAC ACGGAGGCG AGCTTCTTC GCCTGCCACC AGGAGTGACT
```

FIG. 5-9



pCMV-deINS35

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D
5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGCGGGACA
TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGCTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

+2 N T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTT CTGGCTGCC CCGGACTCC GACGTGAGT CCTATTCTC CATGCCCCCC
TATGCTGTTG TAGGAGACTC GGGCGGGAA GACCGACGGG GGGGCTGAG CTGCGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V
BamHI

5441 CTGGAGGGG AGCCTCGCA TCCGGAICTT AGCGACGGT CATGTC AAC GGTCACTAGT GAGGCCAACG CGGAGGATGT
GACCTCCCC TCGACCCCT AGGCTAGAA TCGTGCCCA GTACCAGTTG CCAGTCATCA CTCGGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGCTGC TCAATGCTT ACTCTGGAC AGGGCACTC GTACCCCGT GCGCGCGGA AGAACAGAAA CTGCCCATCA
GCACACGAG AGTTACAGAA TGAGAACCTG TCCGGGTGAG CAGTGGGCA CGCGGGCCT TCTTGCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCACTAAG CAACTCGTTG CTACGTCACC ACAATTTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGCAGAAG
TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTAAACCA CATAAGTGG TGGAGTGGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
TTTCAGTGA AACTGTCTGA CGTTCAAGAC CTGTGGGTAA TGGTCTGCA TGAGTTCTC CAATTCTGTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGTTGACG CTGACGCCCC CACACTCAGC CAAATCCAAG TTGCTTATG
TCACCTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

FIG. 5-10



pCMV-deINS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGCAAAAGA CGTCGGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
CCCCTTTTCT GCAGGCAACG GTACGGTCTT TCCGGGATTG GGTGTAGTG AGGCACACCT TTCTGGAAGA CCTTCTGTIA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGGCT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC
CATTTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCCAG CATTCCGTCG

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCGTCTCATC GTGTTCCTCG ATCTGGGCGT GCGCGTGTGC GAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT
AGCAGAGTAG CACAAGGGGC TAGACCCGCA GCGGCACAG CTTTCTTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGCACAGGG GTTGAATTC TCGTCAAGC GTGGAAGTCC
ACGGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAG AGCACGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAAACCC CAATGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC ACTCACTGAG AGCGACATCC GTACGGGAGGA
TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGGGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCCTCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGCAATCTAC CAATGTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGCG CACCGGTAGT TCAGGAGATG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT CCGGCTATCG CAGGTGCCG GCGAGCGGG TACTGACAACT TAGCTGTGGT
CGGGAGAATG GTTAAGTTCC CCCCCTCTGA CGCCGATAGC GTCCACGGCG CGCTCGCCG ATGACTGTG ATCGACACCA

FIG. 5-11



pCMV-delINS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G
6401 AACACCCCTCA CTTGCTACAT CAAGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTTG AGTGGTACG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
GCTGCTGAAT CAGCAATAGA CACTTTCGGC CCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCGGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
6561 CCAGGTACTC CGCCCCCCT GGGGACCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTC
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641 TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTA CCCTACAACC CCCCTCGGA GAGTGCCTG
AGTCAGCGG TGCTGCGCG ACCTTCTCC CAGATGATGG AGTGGCACT GGGATGTTGG GGGAGCGCT CTCGACGCAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTGCCCC CACACTGTGG GCGAGGATGA
CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
6801 TACTGATGAC CCATTTCCTT AGCGTCCTTA TAGCCAGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
ATGACTACTG GGTAAGAAA TCGCAGGAAT ATCGGTCCCT GGTGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881 TGCTACTCCA TAGAACCCT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
ACGATGAGT ATCTTGGTGA CCTAGATGGA GGTAGTAAG TTTCTGAGT ACCGGAGTCG CGTAAAGTG AGCTGCTAAT

FIG. 5-12



pCMV-delINS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
6961 CTCTCCAGGT GAAATCAATA GGTGGCCGC ATGCCTCAGA AAACCTGGGG TACGCCCTT GCGAGCTTGG AGACACCGGG
GAGAGTCCA CTTTAGTTAT CCCACCGGG TACGGAGTCT TTTGAACCCC ATGGCGGGA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
7041 CCCGAGCGT CCGGCTAGG CTTCTGGCCA GAGGAGCAG GGTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
GGGCTCGCA GCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGGCTGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
TCTGTGTTG AGTTGAGTG AGTTATGCG CGGCGACCG TCGACCTGAA CAGGCCGACC AAGTCCGAC CGATGTGCCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCCCG TGGATCTGGT TTTGCTTACT CCTGCTTGGT GCAGGGGTAG
CCCTCTGTAA ATAGTGTGCG ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAATCTAG AAAGGCGCGC
CGTAGATGGA GGAGGGGTTG CTTACTTCCA ACCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTITAGATC TTTCCGCGCG

BamHI MluI

7361 CAAGATATCA AGGATCCACT ACGGTTAGA GCTCGCTGAT CAGGCTGAC TGTGCTTCT AGTTGCCAGC CATCTGTTGT
GTTCTATAGT TCCTAGGTGA TCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTGG GTAGACAACA

7441 TTGCCCCCTCC CCCGTGCCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCTTA ATAAATGAG GAAATTGCAT
AACGGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA

FIG. 5-13

pCMV-deINS35

7521	CGCATTGTCT GAGTAGGTGT CATTCATTTC TGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT GGCTAACAGA CTCATCCACA GTAAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCTGTTCC CCCTCCTAAC CCTTCTGTTA
7601	AGCAGGCATG CTGGGGAGCT CTTCGGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTGT TCGGCTGCGG CGAGCGGTAT TCGTCCGTAC GACCCCTCGA GAAAGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA
7681	CAGCTCACTC AAAGGCGGTA ATAGGGTTAT CCACAGATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC
7761	CAAAAGGCCA GGAACCGTAA AAAGGCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA GTTTTCGGT CCTTGGCATT TTTCCGGCG AACGACCGCA AAAAGGTATC CGAGGCGGGG GGAAGTCTCG TAGTGTTTT
7841	TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGTAT ATTTCTATGG TCCGCAAAAG GGGACCTTCG AGGAGCACG
7921	GCTCTCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCTTCTC CTTCGGGAA GCGTGGCGCT TTCTCAATGC CGAGAGGACA AGGCTGGGAC GCGGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG
8001	TCACGCTGTA GGTATCTCAG TTCGCTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCGCG TTCAGCCCCG AGTCCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT
8081	CCGCTCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GGCGACGCG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC
8161	GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTAG CTACACTAGA CATGTGCTTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

FIG. 5-14



pCMV-delNS35

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA
TCCTGTGATA AACCATAGAG GCGAGAGCAG TTGGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT

8321 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTCTGC GTCTAATGCG CGTCTTTTTC TCCTAGAGTT CTCTAGGAA

8401 TGATCTTTTC TAGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTITGG TCATGAGATT ATCAAAAAGG
ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

8481 ATCTTCACCT AGATCCTTTT AAATTAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
TAGAAGTGA TCTAGGAAA TTTAATTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCAATTGAA CCAGACTGTC

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG
AATGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGTA TCAACGGACT GAGGGGCAGC

8641 TGTAGATAAC TAGGATACGG GAGGGCTTAC CATCTGGCC CAGTCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACGGG GTACACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCIGCA ACTTTATCCG CCTCCATCCA
GGTCTAAATA GTCGTTATTT GGTCCGTCGG CCTTCCCGGC TCGGCTCTC ACCAGGACGT TGAATAGGC GGAGGTAGGT

8801 GTCTATTAAAT TGTGCCCCG AAGCTAGAGT AAGTAGTTTG CCAGTTAATA GTTTGGGCAA CGTTGTTGCC ATTGCTACAG
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGGCTT GCAACAACGG TAACGATGTC

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGCTTCAAT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
CGTAGCACCA CAGTGGGAGC AGCAAAACCAT ACCGAAAGTAA GTCGAGGCCA AGGTTTGCTA GTTCCGCTCA ATGTACTAGG

FIG. 5-15



pCMV-deINS35

8961	CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTGGTCTCT CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCAATC AACCGCGTC ACAATAGTGA
9041	CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCTATG CCATCCGTA GATGCTTTTC TGTGACTGGT GAGTACTCAA GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT
9121	CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCCTGCCCC GCGTCAATAC GGGATAATAC CGCGCCACAT GGTTCAGTAA GACTCTTATC ACATACGCGC CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA
9201	AGCAGAACTT TAAAAGTGCT CATCAATTGA AACGTTCTT CGGGCGGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTGTGAAGAA GCCCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG
9281	CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCCAG CGTTTCTGGG TGAGCAAAAA GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAACT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT
9361	CAGGAAGGCA AAATGCCGCA AAAAAGGGA TAAGGGGCAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT GTCCTTCCGT TTTACGGCGT TTTTTCCTT ATTCCCGCTG TGCCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA
9441	TATTGAAGCA TTTATCAGG TTATTGCTC ATGAGCGGAT ACATATTGA ATGTATTAG AAAAAATAAC AAATAGGGGT ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTG TTTATCCCCA
9521	TCCGGGCACA TTTCCCCGAA AAGTCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAAATAGG AGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGCTAAT AATAGTACTG TAATTGGATA TTTTATCCG
9601	GTATCACGAG GCCCTTTCT C CATAGTGCTC CGGAAAGCA G

FIG. 5-16



pCMV-II

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAAAGGGAT
AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGGGGGTG TTGGCGGGTG TCGGGGGCTGG CTTAACATG CGGCATCAGA
CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACGCCAC AGCCCGGACC GAATTGATAC GCCGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTCA AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG
CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGTTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGCCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGCGGGA
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATT ATTTTTTTA ATCAGTCGT ACCCGGCTC TTACCGCGCT

321 ACTGGCGGG GAGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
TGACCGGCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCAATTAGT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGGCCCAT
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGTGTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT
CTGCAGTTAT TACTGCATAC AAGGTATCA TTGGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AAAGTGGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACGG TAAATGGCCC
TTTGACGGGT GAACCGTCAT GTAGTTTACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTTACTGCC ATTTACCGGG

FIG. 7-1

pCMV-II

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
CGGACCGTAA TACGGGTCAAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCGAAGT CTCCACCCCA
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATGCCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAAACCCCG CCCGTTGACG
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC
GGTAGGTGG ACAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGG AGGCGCCGC CCTTGCCACG TAACCTTGGC

1121 GGATTCGCC TGCCAAGAGT GACGTAAGTA CGGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGGC ACGGTTCTCA CTGCATTCTAT GCGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATA TTGACCACTC CCCTATTGGT GACGATACCT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCATTTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGG GTCCATTAT
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA

FIG. 7-2



pCMV-II

1441	TATTTACAAA	TTACATATA	CAACAACGCC	GTCCCCCGTG	CCGCGAGTTT	TTATTAAACA	TAGCGTGGGA	TCTCCGACAT
	ATAAATGTTT	AAGTGTATAT	GTTGTGCGG	CAGGGGGCAC	GGCGTCAAA	AATAATTTGT	ATCGCACCT	AGAGGCTGTA
<hr/>								
1521	CTCGGGTACG	TGTTCCGGAC	ATGGGCTCTT	CTCCGGTAGC	GGCGGAGCTT	CCACATCCGA	GCCCTGGTCC	CATCCGTCCA
	GAGCCCATGC	ACAAGGCCCTG	TACCCGAGAA	GAGGCCATCG	CCGCCCTCGAA	GGTGTAGGCT	CGGACCAGG	GTAGGCAGGT
<hr/>								
1601	GCGGCTCATG	GTGGCTCGGC	AGCTCCTTGC	TCCTAACAGT	GGAGGCCAGA	CTTAGCCACA	GCACAATGCC	CACCACCACC
	CGCCGAGTAC	CAGCGAGCCG	TCGAGGAACG	AGGATTGTCA	CCTCCGGTCT	GAATCCGTGT	CGTGTTACGG	GTGGTGGTGG
<hr/>								
1681	AGTGTCCCG	ACAAGGCCGT	GGCGGTAGGG	TATGTGCTCTG	AAAATGAGCT	CGGAGATTGG	GCTCGCACCT	GGACGCAGAT
	TCACACGGCG	TGTTCCGGCA	CGCCATCCC	ATACACAGAC	TTTTACTCGA	GCCTCTAACC	CGAGCGTGGG	CCTGCGTCTA
<hr/>								
1761	GGAAGACTTA	AGCAGCGCG	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTGT	ATTCTGATAA	GAGTCAAGGG	TAACTCCCGT
	CCTTCTGAAT	TCCGTGCGCG	TCTTCTTCTA	CGTCCGTCGA	CTCAACAACA	TAAGACTATT	CTCAGTCTCC	ATTGAGGGCA
<hr/>								
1841	TGCGGTGCTG	TTAACGGTGG	AGGCGAGTGT	AGTCTCAGCA	GTACTCGTTG	CTGCCGGCGG	CGCCACCAGA	CATAATAGCT
	ACGCCACGAC	AATTGCCACC	TCCCGTCACA	TCAGACTCGT	CATGAGCAAC	GACGGCGCGG	GCGGTGGTCT	GTATTATCGA
<hr/>								
EcoRI -----								
1921	GACAGACTAA	CAGACTGTTT	CTTTCCATGG	GTCFTTTCTG	CAGTCACCGT	CGTCGACCTA	AGAATTCAGA	CTCGAGCAAG
	CTGTCTGATT	GTCTGACAA	GAAAGGTACC	CAGAAAAGAC	GTCAGTGGCA	GCAGCTGGAT	TCTTAACTCT	GAGCTCGTTC
<hr/>								
XbaI BamHI MluI -----								
2001	TCTAGAAAGG	CGCGCCAAGA	TATCAAGCAT	CCACTACGGC	TTAGAGCTCG	CTGATCAGCC	TGCAGTGTGC	CTTCTAGTTG
	AGATCTTTCC	GCGCGGTTCT	ATAGTTCTTA	GGTGATGGCG	AATCTCGAGC	GACTAGTCGG	AGCTGACACG	GAAGATCAAC

FIG. 7-3



pCMV-II

2081 CCAGGCATCT GTTGTGTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCTCTT TCCTAATAAA
GGTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGGATTATTT

2161 ATGAGGAAAT TGCATCGCAT TGCTGAGTA GGTGTCATT TATTCTGGG GGTGGGTGG GGCAGGACAG CAAGGGGGAG
TACTCCTTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCC CCACCCCAACC CCGTCTCTGTC GTTCCCCCTC

2241 GATTGGGAAG ACAATAGCAG GCATGCTGG GAGCTCTTCC GCTTCCTCGC TCACCTGACTC GCTGGGCTCG GTCTTCTGGC
CTAACCCCTC TGTATCGTC CTAACGACCC CTCGAGAAG CGAAGGAGCG AGTGACTGAG CGACGGGAGC CAGCAAGCCC

2321 TCGGGCGAGC GGTATCAGCT CACTCAAAG CGGTAATACG GTTATCCACA GAATCAGGG ATAACGCAGG AAAGAACATG
ACGCCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTTGTA

2401 TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAG CCGCGTTGCT GCGGTTTTTC CATAGGCTCC GCCCCCTGA
ACTCGTTTTT CCGTCGTTTT CCGGTCCTTG GCATTTTTTC GCGCAACGA CCGCAAAAAG GTATCCGAG CGGGGGGACT

2481 CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGCG TTTCCCGCTG
GCTCGTAGTG TTTTITAGCTG CGAGTTCAGT CTCACCGCT TTGGGCTGTC CTGATAATTC TATGGTCCGC AAAGGGGAC

2561 GAAGCTCCCT CGTGGGCTCT CCTGTTCGA CCTGCGGCT TACCGGATAC CTGTCCGCTT TTTCTCCCTC GGGAAGCGTG
CTTCGAGGGA GCACGGGAGA GAACAAGGCT GGGACGGCGA ATGGCTATG GACAGGCGGA AAGAGGGAAG CCTTTCGCAC

2641 GCGCTTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCCG TGTAGGTGCT TCGGTCCAAG CTGGGCTGTG TGCACGAACC
CGCGAAAGAG TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA ACCGAGCTTC GACCCGACAC ACGTGTCTGG

2721 CCCCCCTCAG CCCGACCGCT GCGCCTTATC CGGTAACAT CGTCTTGAGT CCAACCCGCT AAGACACGAC TTATCGCCAC
GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GGTGGGGCA TTCTGTGCTG AATAGCGGTG

FIG. 7-4

pCMV-II

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
ACCGTCCGCG GTGACCAATTG TCCTAATCGT CTGCTCCAT ACATCGGCCA CGATGTCTCA AGAATTCAC CACCGGATTG

2881 TAGGGCTACA CTAGAAAGGAC AGTATTGGT ATCTGGGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC
ATGCCGATGT GATCTTCCTG TCATAAACCA TAGACGGGAG ACGACTTCGG TCAATGGAAG CCTTTTCTC AACCATCGAG

2961 TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTT TTGTTTGCA AGCAGCAGAT TACGGGCAGA AAAAAAGGAT
AACTAGGCCG TTTGTTTGGT GCGGACCATC GCCACCAAAA AAACAAACGT TCGTCTCTA ATGCGCGTCT TTTTTCCTA

3041 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTIGAGTG CAATTCCTA AAACCACTAC

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAAT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
TCTAATAGTT TTTCTAGAA GTGGATCTAG GAAAAATTAA TTTTACTTC AAAATTAGT TAGATTTTAT ATATACTCAT

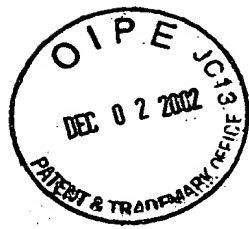
3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCTGTCA TCCATAGTTG
TTGAACCCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

3281 CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT GGCCCCCAGTG CTGCAATGAT ACCGCGAGAC
GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTAC GACGTIAC TA TGGCGCTCTG

3361 CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTCT CTGCAACTTT
GGTGGAGTG GCCGAGGTCT AAATAGTCTG TATTGGTCTG GTCCGCTTC CCGGCTCGG TCTTCACCAG GACGTTGAAA

3441 ATCCGCCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
TAGGCGGAGG TAGTTCAGAT AATTACAAC GGCCCTTGA TCTCATTTAT CAAGCGGTCA ATTATCAAAAC GCGTTGCAAC

FIG. 7-5



pCMV-II

3521	TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG AACGGTAACG ATGTCGCTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGCT TGCTAGTTCC
3601	CGAGTTACAT GATCCCCCAT GTTGTCGAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTGAGAA GTAAGTTGGC GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCGCCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCACCCG
3681	CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA GCGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT
3761	CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGCGACCG AGTTGCTCTT GCCGGCGGC AATACGGGAT GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCGCTGGC TCAACGAGAA CGGGCCGCG TTTATGCCCTA
3841	AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAGACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GC'TTTTGAGA GTTCCTAGAA
3921	ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCAGCT GGGTTGACTA GAACTCGTAG AAAATGAAAG TGGTCGCCAA
4001	CTGGGTGAGC AAAAAACAGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC GACCCACTCG TTTTGTCTT TCCGTTTTC TCCCTATTCC CGCTGTGCCT TTACAACCTTA TGAGTATGAG
4081	TTCCCTTTTTC AATATTATTG AAGCATTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA AAGGAAAAAG TTATAATAAC TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTT
4161	TAAACAAATA GGGGTTCGCG GCACATTTC CCGAAAAAGTG CCACCTGAG TCTAAGAAAC CATTATTATC ATGACATTAA ATTTGTTTAT CCCCAAGGCG CGTGTAAGG GGC'TTTTCAC GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT
4241	CCTATAAAAA TAGCGGTATC ACGAGGCCCT TTCGTC GGATATTTT ATCCGCATAG TGCTCCGGGA AAGCAG

FIG. 7-6

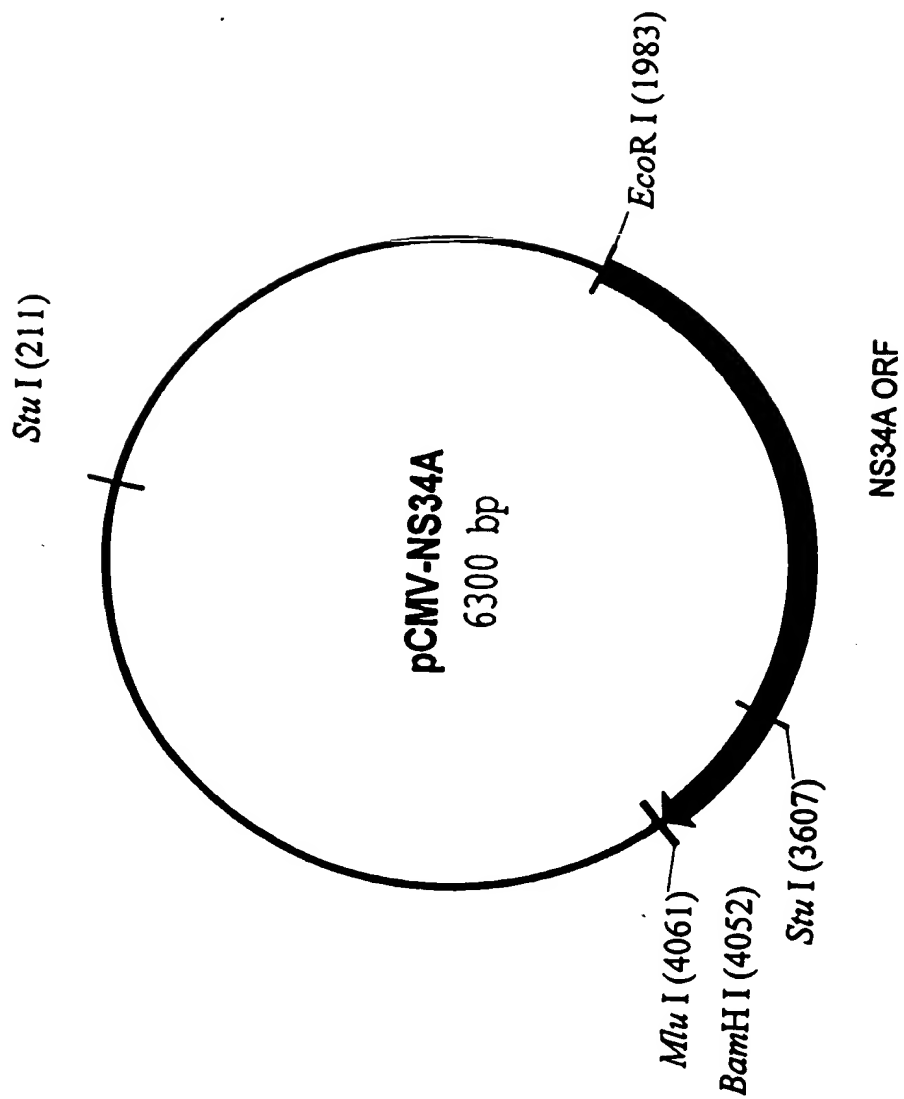


FIG. 8



pCMV-NS34A

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC
51	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT	GCCGGGAGCA	GACAAGCCCG
	CTCTGCCAGT	GTCGAACAGA	CATTGCGCTA	CGGCCCTCGT	CTGTTGCGGC
101	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATG
	AGTCCCGCGC	AGTCGCCAC	AACCGCCAC	AGCCCCGACC	GAATTGATAC
151	CGGCATCAGA	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTGCA
	GCCGTAGTGT	CGTCTAACAT	GACTCTCACG	TGGTATACTT	CGAAAAACGT
StuI					
201	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG	AATAGCTCAG
	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC	TTATCGAGTC
251	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAAT	TAGTCAGCCA
	TCCGGCTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT
301	TGGGGCGGAG	AATGGGCGGA	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG
	ACCCCGCCTC	TTACCCGCCT	TGACCCGCC	CTCCCTTAAT	AACCGATAAC
351	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA
	GTACAGGTTA	TACTGGCGGT	ACAACTGTAA	CTAATAACTG	ATCAATAATT
451	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGCCCATATA	TGGAGTTCGG
	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA	TCGGGTATAT	ACCTCAAGGC
501	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC
	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTTGCTGG
551	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA
	GGGCGGGTAA	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT
601	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AAACTGCCCA
	CCCTGAAAGG	TAACTGCAGT	TACCCACCTC	ATAAATGCCA	TTTGACGGGT
651	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCC	CCTATTGACG
	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC
701	TCAATGACGG	TAAATGGCCC	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA
	AGTTACTGCC	ATTTACCGGG	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT
751	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC
851	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG
	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT	AACTGCAGTT	ACCTCAAAC

FIG. 9-1



pCMV-NS34A

901	TTTTGGCACC AAAACCGTGG	AAAATCAACG TTTGTAGTTC	GGACTTTCCA CCTGAAAGGT	AAATGTCGTA TTTACAGCAT	ATAACCCCGC TATTGGGGCG
951	CCCGTTGACG GGGCAACTGC	CAAATGGGCG GTTTACCCGC	GTAGGCGTGT CATCCGCACA	ACGGTGGGAG TGCCACCCTC	GTCTATATAA CAGATATATT
1001	GCAGAGCTCG CGTCTCGAGC	TTTAGTGAAC AAATCACTTG	CGTCAGATCG GCAGTCTAGC	CCTGGAGACG GGACCTCTGC	CCATCCACGC GGTAGGTGCG
1051	TGTTTTGACC ACAAAACCTG	TCCATAGAAG AGGTATCTTC	ACACCGGGAC TGTGGCCCTG	CGATCCAGCC GCTAGGTCGG	TCCGCGGCCG AGGCGCCCGC
1101	GGAACGGTGC CCTTGCCACG	ATTGGAACGC TAACCTTGCG	GGATTCCCCG CCTAAGGGGC	TGCCAAGAGT ACGGTTCTCA	GACGTAAGTA CTGCATTTCAT
1151	CCGCCTATAG GGCGGATATC	ACTCTATAGG TGAGATATCC	CACACCCCTT GTGTGGGGAA	TGGCTCTTAT ACCGAGAATA	GCATGCTATA CGTACGATAT
1201	CTGTTTTTGG GACAAAAC	CTTGGGGCCT GAACCCCGGA	ATACACCCCC TATGTGGGGG	GCTCCTTATG CGAGGAATAC	CTATAGGTGA GATATCCACT
1251	TGGTATAGCT ACCATATCGA	TAGCCTATAG ATCGGATATC	GTGTGGGTTA CACACCCAAT	TTGACCATTA AACTGGTAAT	TTGACCACTC AACTGGTGAG
1301	CCCTATTGGT GGGATAACCA	GACGATACTT CTGCTATGAA	TCCATTACTA AGGTAATGAT	ATCCATAACA TAGGTATTGT	TGGCTCTTTG ACCGAGAAAC
1351	CCACAACATAT GGTGTGATA	CTCTATTGGC GAGATAACCG	TATATGCCAA ATATACGGTT	TACTCTGTCC ATGAGACAGG	TTCAGAGACT AAGTCTCTGA
1401	GACACGGACT CTGTGCCTGA	CTGTATTTTT GACATAAAAA	ACAGGATGGG TGTCCTACCC	GTCCATTTAT CAGGTAAATA	TATTTACAAA ATAAATGTTT
1451	TTCACATATA AAGTGTATAT	CAACAACGCC GTTGTTGCCG	GTCCCCCGTG CAGGGGGCAC	CCCGCAGTTT GGGCGTCAA	TTATTAAACA AATAATTTGT
1501	TAGCGTGGA ATCGCACCT	TCTCCGACAT AGAGGCTGTA	CTCGGGTACG GAGCCCATGC	TGTTCCGGAC ACAAGGCCTG	ATGGGCTCTT TACCCGAGAA
1551	CTCCGGTAGC GAGGCCATCG	GGCGGAGCTT CCGCCTCGAA	CCACATCCGA GGTGTAGGCT	GCCCTGGTCC CGGGACCAGG	CATCCGTCCA GTAGGCAGGT
1601	GCGGCTCATG CGCCGAGTAC	GTCGCTCGGC CAGCGAGCCG	AGCTCCTTGC TCGAGGAACG	TCCTAACAGT AGGATTGTCA	GGAGGCCAGA CCTCCGGTCT
1651	CTTAGGCACA GAATCCGTGT	GCACAATGCC CGTGTTACGG	CACCACCACC GTGGTGGTGG	AGTGTGCCGC TCACACGGCG	ACAAGGCCGT TGTTCCGGCA
1701	GGCGGTAGGG CCGCCATCCC	TATGTGTCTG ATACACAGAC	AAAATGAGCT TTTTACTCGA	CGGAGATTGG GCCTCTAACC	GCTCGCACCT CGAGCGTGGA
1751	GGACGCAGAT CCTGCGTCTA	GGAAGACTTA CCTTCTGAAT	AGGCAGCGGC TCCGTCGCCG	AGAAGAAGAT TCTTCTTCTA	GCAGGCAGCT CGTCCGTCGA
1801	GAGTTGTTGT CTCAACAACA	ATTCTGATAA TAAGACTATT	GAGTCAGAGG CTCAGTCTCC	TAATCCCGT ATTGAGGGCA	TGCGGTGCTG ACGCCACGAC

FIG. 9-2



pCMV-NS34A

```

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
    AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
    GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2                                     M A P
                                     EcoRI
                                     ~~~~~
1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCGCCCA
    CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

+2 I T A Y A Q Q T R G L L G C I I T
2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC
    AGTGCCGCAT GCGGGTCGTC TGTTCCTCCG AGGATCCAC GTATTAGTGG

+2 S L T G R D K N Q V E G E V Q I V
2051 AGCCTAACTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT
    TCGGATTGAC CGGCCCTGTT TTTGGTTCAC CTCCCACTCC AGGTCTAACA

+2 S T A A Q T F L A T C I N G V C
2101 GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT
    CAGTTGAEGA CGGGTTTGA AGGACCGTTG CACGTAGTTA CCCACACGA

+2 W T V Y H G A G T R T I A S P K G
2151 GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT
    CCTGACAGAT GGTGCCCCGG CTTTGCTCCT GGTAGCGCAG TGGGTTCCTA

-2 P V I Q M Y T N V D Q D L V G W P
2201 CCTGTCTATCC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC
    GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

+2 A S Q G T R S L T P C T C G S S
2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG
    GCGAAGCGTT CCATGGGCGA GTAAGTGTGG GACGTGAACG CCGAGGAGCC

+2 D L Y L V T R H A D V I P V R R R
2301 ACCTTTACCT GGTCACGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG
    TGGAAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC

+2 G D S R G S L L S P R P I S Y L K
2351 GGTGATAGCA GGGGCAGCCT GCTGTCGCCC CGGCCCATTT CCTACTTGAA
    CCACTATCGT CCCCCTCGGA CGACAGCGGG GCCGGGTAAA GGATGAACTT

+2 G S S G G P L L C P A G H A V G
2401 AGGCTCCTCG GGGGGTCCGC TGTTGTGCCC CGCGGGGCAC GCCGTGGGCA
    TCCGAGGAGC CCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT

+2 I F R A A V C T R G V A K A V D F
2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT
    ATAAATCCCG GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA

+2 I P V E N L E T T M R S P V F T D
2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTACGGA
    TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCTT
  
```

FIG. 9-3



pcMV-NS34A

+2 N S S P P V V P Q S F Q V A H L
2551 TAACTCCTCT CCACCACTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC
ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

+2 H A P T G S G K S T K V P A A Y A
2601 ATGCTCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA
TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT

+2 A Q G Y K V L V L N P S V A A T L
2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT
CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

+2 G F G A Y M S K A H G I D P N I
2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA
CCCCAAACCA CGAATGTACA GGTTCGAGT ACCCTAGCTA GGATTGTAGT

+2 R T G V R T I T T G S P I T Y S T
2751 GGACCGGGGT GAGAACAATT ACCACTGGCA GCCCCATCAC GTACTCCACC
CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG

+2 Y G K F L A D G G C S G G A Y D I
2801 TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT
ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCGC GAATACTGTA

+2 I I C D E C H S T D A T S I L G
2851 AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA
TTATTAAACA CTGCTCACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT

+2 I G T V L D Q A E T A G A R L V V
2901 TTGGCACTGT CTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG
AACCGTGACA GGAAGTGGT CGTCTCTGAC GCCCCGCTC TGACCAACAC

+2 L A T A T P P G S V T V P H P N I
2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT
GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA

+2 E E V A L S T T G E I P F Y G K
3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG
GTCCTCCAA CGAGACAGGT GGTGGCCTCT CTAGGGAAAA ATGCCGTTCC

+2 A I P L E V I K G G R H L I F C H
3051 CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTTCTGTGAT
GATAGGGGGA GCTTCATTAG TTCCCCCCT CTGTAGAGTA GAAGACAGTA

+2 S K K K C D E L A A K L V A L G I
3101 TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTCG CATTGGGCAT
AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA

+2 N A V A Y Y R G L D V S V I P T
3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA
GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT

+2 S G D V V V V A T D A L M T G Y T
3201 GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC
CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG

FIG. 9-4



pCMV-NS34A

+2 G D F D S V I D C N T C V T Q T V
3251 GGC GACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT
CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P
3301 CGATTTTCAGC CTTGACCCTA CCTTCACCAT TGAGACAATC ACGCTCCCCC
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCGAGGGGG

+2 Q D A V S R T Q R R G R T G R G K
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG
TTCTACGAGA GAGGGCGTGA GTTGCAGCCC CGTCCTGACC GTCCCCCTC

+2 P G I Y R F V A P G E R P S G M F
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT
GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2 D S S V L C E C Y D A G C A W Y
3451 CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC
TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F
3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT
GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTAAAAACCC TCCCGCAGAA

+2 T G L T H I D A H F L S Q T K Q
StuI
~~~~~  
3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA  
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTTCGTCT

---

+2 S G E N L P Y L V A Y Q A T V C A  
3651 GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT  
CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA

---

+2 R A Q A P P P S W D Q M W K C L I  
3701 AGGGCTCAAG CCCCTCCCC ATCGTGGGAC CAGATGTGGA AGTGTTTGAT  
TCCCGAGTTC GGGGAGGGGG TAGCACCTG GTCTACACCT TCACAAACTA

---

+2 R L K P T L H G P T P L L Y R L  
3751 TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG  
AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC

---

+2 G A V Q N E I T L T H P V T K Y I  
3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC  
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCAGTG GTTTATGTAG

---

+2 M T C M S A D L E V V T S T W V L  
3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT  
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

---

+2 V G G V L A A L A A Y C L S T G  
3901 CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCCTG TCAACAGGCT  
GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA

FIG. 9-5



# pCMV-NS34A

|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
|------|----|-------|-------|------|------|-----|-----|-------|-------|-----|-----|-------|-------|------|------|-------|-------|------|-----|-----|-------|-----|-----|
|      | +2 | C     | V     | V    | I    | V   | G   | R     | V     | V   | L   | S     | G     | K    | P    | A     | I     | I    |     |     |       |     |     |
| 3951 |    | GC    | TG    | GC   | AT   | AG  | TG  | GG    | GC    | AG  | GT  | CG    | TCT   | TGT  | CC   | GG    | GA    | AG   | CC  |     |       |     |     |
|      |    | CG    | CA    | CC   | AG   | TA  | CC  | CC    | GT    | CC  | CA  | GC    | AG    | AA   | CA   | GC    | CC    | TT   | CG  |     |       |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
|      | +2 | P     | D     | R    | E    | V   | L   | Y     | R     | E   | F   | D     | E     | M    | E    | E     | C     |      |     |     |       |     |     |
| 4001 |    | CCT   | GAC   | AGG  | GA   | AAG | TCT | CTA   | CCG   | AG  | AGT | TTC   | GAT   | GAG  | ATG  | G     | AAG   | AGT  | GC  | TA  |       |     |     |
|      |    | GG    | ACT   | GT   | CCC  | TT  | CAG | GAG   | AT    | GG  | CT  | CT    | CA    | AG   | CT   | ACT   | CT    | ACC  | TT  | CT  | CAC   | GC  | AT  |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
|      |    | Bam   | HI    | Mlu  | I    |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
|      |    | ----- | ----- |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4051 |    | GG    | AT    | CC   | ACT  | CG  | CT  | TAG   | AG    | CT  | CG  | CT    | GAT   | C    | AG   | CC    | TC    | GACT | GT  | GC  | CT    | TCT | TA  |
|      |    | CCT   | AG    | GT   | GAT  | GCG | CA  | AT    | CTC   | GAG | CG  | ACT   | AG    |      | TC   | GG    | AG    | CT   | GA  | CAC | GG    | AAG | AT  |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4101 |    | GT    | TGCC  | AG   | CC   | AT  | CT  | GT    | TG    | TT  | TG  | CC    | CC    | TCCC | CC   | GT    | GC    | CT   | TTC | CT  | TG    | ACC | CTG |
|      |    | CA    | AC    | GG   | TC   | GG  | TAG | ACA   | ACAA  | AC  | GG  | GG    | GAG   | GG   | GG   | CA    | CG    | GA   | AG  | GA  | ACT   | TGG | GAC |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4151 |    | GA    | AGGT  | GC   | CA   | CT  | CC  | CA    | CT    | GT  | CT  | TTC   | CT    | TAA  | TAAA | AT    | GAG   | G    | AA  | AT  | TG    | CAT | C   |
|      |    | CT    | TCC   | AC   | GGT  | GAG | GGT | GACA  | GG    | AA  | AGG | ATT   | AT    | TTC  | ACT  | CC    | TT    | TA   | AC  | G   | TAG   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4201 |    | GC    | ATT   | GT   | CTG  | AG  | TAG | GT    | GTC   | ATT | CT  | ATT   | CT    | GG   | GG   | GGT   | GGG   | GT   | GG  | GG  | CAG   | G   |     |
|      |    | CG    | TA    | AC   | AG   | TC  | AT  | CC    | AC    | TA  | AG  | ATA   | AG    | CC   | CC   | CC    | CA    | CC   | CA  | CC  | CC    | CG  | TCC |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4251 |    | AC    | AG    | CA   | AGG  | GG  | AG  | GAT   | TGG   | GA  | AG  | CA    | ATA   | GC   | AG   | GC    | AT    | GC   | TG  | GG  | GAG   | CTC |     |
|      |    | TG    | TC    | GT   | TCCC | CCT | CCT | AA    | ACC   | CT  | TCT | GTT   | AT    | CG   | TCC  | GT    | ACG   | ACC  | CC  | TC  | GC    | AG  |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4301 |    | TT    | CC    | GCT  | TCC  | TC  | GC  | TACT  | G     | ACT | CG  | CT    | GCG   | CT   | CG   | GT    | CG    | TT   | CG  | GT  | GC    | GGC |     |
|      |    | AAG   | GC    | GA   | AGG  | AG  | CG  | AG    | TG    | AG  | CG  | CG    | AC    | GC   | AG   | CC    | AG    | CA   | GC  | CG  | AC    | CGC |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4351 |    | GAG   | CG    | G    | TATC | AG  | CT  | CA    | CTCA  | AAG | GC  | GG    | TAA   | TAC  | GG   | T     | TATC  | CAC  | AG  | AAT | CA    | TA  |     |
|      |    | CT    | CG    | CA   | TAG  | TC  | GAG | T     | GAGT  | TT  | CC  | GC    | CATT  | AT   | G    | CC    | AATAG | GT   | G   | TCT | T     | AGT |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4401 |    | GG    | GG    | GATA | AACG | CAG | GAA | AAGAA | CAT   | GT  | GAG | CA    | AA    | AG   | GC   | CAGC  | AAA   | AG   | GC  | CC  | AG    |     |     |
|      |    | CCC   | CT    | ATT  | TGC  | GTC | CTT | TCTT  | GT    | AC  | ACT | CGT   | TTT   | CC   | GG   | TCG   | TTT   | CC   | GG  | TC  |       |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4451 |    | GA    | ACC   | G    | TAAA | AAG | GC  | CG    | CGT   | TG  | CT  | GG    | CGT   | TTT  | CC   | AT    | AGG   | CT   | CC  | GC  | CCCC  |     |     |
|      |    | CT    | TG    | GC   | ATT  | TT  | CC  | GG    | CGCA  | AC  | G   | ACC   | CGCA  | AA   | AG   | G     | TATCC | GAG  | GC  | GG  | GGG   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4501 |    | CT    | GAC   | GAG  | CA   | TC  | AC  | AAAA  | AAT   | CG  | AC  | GCT   | CAA   | GT   | CAG  | AGG   | TG    | GCG  | AA  | AC  | CCG   |     |     |
|      |    | GAC   | TG    | CT   | CGT  | AG  | TG  | TTTT  | TTA   | GCT | G   | CG    | AGT   | CAG  | T    | CT    | CCAC  | CG   | CTT | TG  | GGC   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4551 |    | AC    | AGG   | ACT  | TAT  | AA  | AG  | A     | TACCA | GG  | CG  | TTT   | TCCC  | CCT  | GGA  | AG    | CT    | CC   | CT  | CG  | TGCG  |     |     |
|      |    | TG    | T     | CT   | GATA | TTT | CT  | AT    | TGGT  | CC  | G   | CAA   | AGGG  | GG   | AC   | CT    | TCGA  | GG   | GAG | CAC | GC    |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4601 |    | CT    | CT    | CCT  | GTT  | CC  | GAC | CC    | TGC   | CG  | CT  | TAC   | CGG   | AT   | AC   | CT    | GTCC  | GC   | CTT | TCT | CC    |     |     |
|      |    | GAG   | AGG   | GAC  | AA   | GG  | CT  | GGG   | GACG  | GCG | AAT | TGG   | CC    | TAT  | GG   | AC    | AGG   | CG   | GAA | AG  | AGG   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4651 |    | CT    | T     | CGG  | GAAG | CG  | TGG | CGCT  | T     | TCT | CA  | AT    | GCT   | CAC  | GCT  | G     | TAG   | GT   | AT  | CT  | CAGT  |     |     |
|      |    | GA    | AG    | CC   | CTTC | GC  | ACC | CGCA  | AA    | AG  | ATT | ACGA  | GT    | G    | CG   | ACATC | CAT   | AG   | A   | G   | TCA   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4701 |    | TC    | GGT   | G    | TAGG | TC  | GTT | CGCT  | C     | CA  | AG  | TGGG  | C     | TGT  | G    | TG    | CACG  | A    | AC  | CC  | CCCGT |     |     |
|      |    | AG    | CC    | AC   | ATCC | AG  | CA  | AG    | CGAG  | GT  | T   | CG    | ACCCG | AC   | AC   | AC    | GTGC  | TT   | G   | G   | G     | G   | G   |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4751 |    | TC    | AG    | CC   | CGAC | CG  | CT  | GCG   | CCT   | TAT | CC  | GG    | TAA   | CT   | AT   | CG    | TCTT  | GAG  | T   | CA  | AA    | CC  |     |
|      |    | AG    | T     | CGG  | CTG  | GCG | AC  | GCGGA | AA    | AT  | AGG | CCATT | GAT   | AG   | CAG  | AA    | CT    | CAG  | G   | T   | TGG   |     |     |
|      |    |       |       |      |      |     |     |       |       |     |     |       |       |      |      |       |       |      |     |     |       |     |     |
| 4801 |    | CG    | GTA   | AG   | GACA | CG  | ACT | TAT   | CG    | CC  | ACT | G     | GCAG  | CAG  | CC   | ACT   | G     | TA   | AC  | AG  | GATT  |     |     |
|      |    | G     | CC    | ATT  | CTGT | G   | CT  | GA    | ATAGC | GG  | T   | GA    | CCGTC | G    | T    | CGG   | TGACC | ATT  | G   | T   | CTTAA |     |     |

FIG. 9-6



pCMV-NS34A

|      |             |            |             |            |            |
|------|-------------|------------|-------------|------------|------------|
| 4851 | AGCAGAGCGA  | GGTATGTAGG | CGGTGCTACA  | GAGTTCTTGA | AGTGGTGGCC |
|      | TCGTCTCGCT  | CCATACATCC | GCCACGATGT  | CTCAAGAACT | TCACCACCGG |
| 4901 | TAACTACGGC  | TACACTAGAA | GGACAGTATT  | TGGTATCTGC | GCTCTGCTGA |
|      | ATTGATGCCG  | ATGTGATCTT | CCTGTCATAA  | ACCATAGACG | CGAGACGACT |
| 4951 | AGCCAGTTAC  | CTTCGGAAAA | AGAGTTGGTA  | GCTCTTGATC | CGGCAAACAA |
|      | TCGGTCAATG  | GAAGCCTTTT | TCTCAACCAT  | CGAGAACTAG | GCCGTTTGTT |
| 5001 | ACCACCGCTG  | GTAGCGGTGG | TTTTTTTGTT  | TGCAAGCAGC | AGATTACGCG |
|      | TGGTGGCGAC  | CATCGCCACC | AAAAAAACAA  | ACGTTTCGTC | TCTAATGCGC |
| 5051 | CAGAAAAAAA  | GGATCTCAAG | AAGATCCTTT  | GATCTTTTCT | ACGGGGTCTG |
|      | GTCTTTTTTT  | CCTAGAGTTC | TTCTAGGAAA  | CTAGAAAAGA | TGCCCCAGAC |
| 5101 | ACGCTCAGTG  | GAACGAAAAC | TCACGTAAAG  | GGATTTTGGT | CATGAGATTA |
|      | TGCGAGTCAC  | CTTGCTTTTG | AGTGCAATTC  | CCTAAAACCA | GTACTCTAAT |
| 5151 | TCAAAAAGGA  | TCTTCACCTA | GATCCTTTTA  | AATTAAAAAT | GAAGTTTAA  |
|      | AGTTTTTCCT  | AGAAGTGGAT | CTAGGAAAAAT | TTAATTTTAA | CTTCAAAATT |
| 5201 | ATCAATCTAA  | AGTATATATG | AGTAAACTTG  | GTCTGACAGT | TACCAATGCT |
|      | TAGTTAGATT  | TCATATATAC | TCATTGAAAC  | CAGACTGTCA | ATGGTTACGA |
| 5251 | TAATCAGTGA  | GGCACCTATC | TCAGCGATCT  | GTCTATTTTC | TTCATCCATA |
|      | ATTAGTCACT  | CCGTGGATAG | AGTCGCTAGA  | CAGATAAAGC | AAGTAGGTAT |
| 5301 | GTTGCCTGAC  | TCCCCGTCGT | GTAGATAACT  | ACGATACGGG | AGGGCTTACC |
|      | CAACGGACTG  | AGGGGCAGCA | CATCTATTGA  | TGCTATGCCC | TCCCGAATGG |
| 5351 | ATCTGGCCCC  | AGTGCTGCAA | TGATACCGCG  | AGACCCACGC | TCACCGGCTC |
|      | TAGACCGGGG  | TCACGACGTT | ACTATGGCGC  | TCTGGGTGCG | AGTGGCCGAG |
| 5401 | CAGATTTATC  | AGCAATAAAC | CAGCCAGCCG  | GAAGGGCCGA | GCGCAGAAGT |
|      | GTCTAAATAG  | TCGTTATTTG | GTCGGTCCGC  | CTTCCCGGCT | CGCGTCTTCA |
| 5451 | GGTCCTGCAA  | CTTTATCCGC | CTCCATCCAG  | TCTATTAATT | GTTGCCGGGA |
|      | CCAGGACGTT  | GAAATAGGCG | GAGGTAGGTC  | AGATAATTAA | CAACGGCCCT |
| 5501 | AGCTAGAGTA  | AGTAGTTCGC | CAGTTAATAG  | TTTGCGCAAC | GTTGTTGCCA |
|      | TCGATCTCAT  | TCATCAAGCG | GTCAATTATC  | AAACGCGTTG | CAACAACGGT |
| 5551 | TTGCTACAGG  | CATCGTGGTG | TCACGCTCGT  | CGTTTGGTAT | GGCTTCATTC |
|      | AACGATGTCC  | GTAGCACCAC | AGTGCGAGCA  | GCAAACCATA | CCGAAGTAAG |
| 5601 | AGCTCCGGTT  | CCCAACGATC | AAGGCGAGTT  | ACATGATCCC | CCATGTTGTG |
|      | TCGAGGCCAA  | GGGTTGCTAG | TTCCGCTCAA  | TGTACTAGGG | GGTACAACAC |
| 5651 | CAAAAAAGCG  | GTTAGCTCCT | TCGGTCCTCC  | GATCGTTGTC | AGAAGTAAGT |
|      | GTTTTTTCGC  | CAATCGAGGA | AGCCAGGAGG  | CTAGCAACAG | TCTTCATTCA |
| 5701 | TGGCCGCGAGT | GTTATCACTC | ATGGTTATGG  | CAGCACTGCA | TAATTCTCTT |
|      | ACCGGCGTCA  | CAATAGTGAG | TACCAATACC  | GTCGTGACGT | ATTAAGAGAA |
| 5751 | ACTGTCATGC  | CATCCGTAAG | ATGCTTTTCT  | GTGACTGGTG | AGTACTCAAC |
|      | TGACAGTACG  | GTAGGCATTC | TACGAAAAGA  | CACTGACCAC | TCATGAGTTG |

FIG. 9-7



pCMV-NS34A

|      |             |            |            |            |            |
|------|-------------|------------|------------|------------|------------|
| 5801 | CAAGTCATTG  | TGAGAATAGT | GTATGCGGCG | ACCGAGTTGC | TCTTGCCCGG |
|      | GTTTCAGTAAG | ACTCTTATCA | CATACGCCGC | TGGCTCAACG | AGAACGGGCG |
| 5851 | CGTCAATACG  | GGATAATACC | GCGCCACATA | GCAGAACTTT | AAAAGTGCTC |
|      | GCAGTTATGC  | CCTATTATGG | CGCGGTGTAT | CGTCTTGAAA | TTTTCACGAG |
| 5901 | ATCATTGGAA  | AACGTTCTTC | GGGGCGAAAA | CTCTCAAGGA | TCTTACCGCT |
|      | TAGTAACCTT  | TTGCAAGAAG | CCCCGCTTTT | GAGAGTTCCT | AGAATGGCGA |
| 5951 | GTTGAGATCC  | AGTTCGATGT | AACCCACTCG | TGCACCCAAC | TGATCTTCAG |
|      | CAACTCTAGG  | TCAAGCTACA | TTGGGTGAGC | ACGTGGGTTG | ACTAGAAGTC |
| 6001 | CATCTTTTAC  | TTTCACCAGC | GTTTCTGGGT | GAGCAAAAAC | AGGAAGGCAA |
|      | GTAGAAAATG  | AAAGTGGTCG | CAAAGACCCA | CTCGTTTTTG | TCCTTCCGTT |
| 6051 | AATGCCGCAA  | AAAAGGGAAT | AAGGGCGACA | CGGAAATGTT | GAATACTCAT |
|      | TTACGGCGTT  | TTTTCCCTTA | TTCCCGCTGT | GCCTTTACAA | CTTATGAGTA |
| 6101 | ACTCTTCCTT  | TTTCAATATT | ATTGAAGCAT | TTATCAGGGT | TATTGTCTCA |
|      | TGAGAAGGAA  | AAAGTTATAA | TAACTTCGTA | AATAGTCCCA | ATAACAGAGT |
| 6151 | TGAGCGGATA  | CATATTTGAA | TGTATTTAGA | AAAATAAACA | AATAGGGGTT |
|      | ACTCGCCTAT  | GTATAAACTT | ACATAAATCT | TTTTATTTGT | TTATCCCCAA |
| 6201 | CCGCGCACAT  | TTCCCCGAAA | AGTGCCACCT | GACGTCTAAG | AAACCATTAT |
|      | GGCGCGTGTA  | AAGGGGCTTT | TCACGGTGGA | CTGCAGATTC | TTTGGTAATA |
| 6251 | TATCATGACA  | TTAACCTATA | AAAATAGGCG | TATCACGAGG | CCCTTTCGTC |
|      | ATAGTACTGT  | AATTGGATAT | TTTTATCCGC | ATAGTGCTCC | GGGAAAGCAG |

FIG. 9-8



Diagram 1

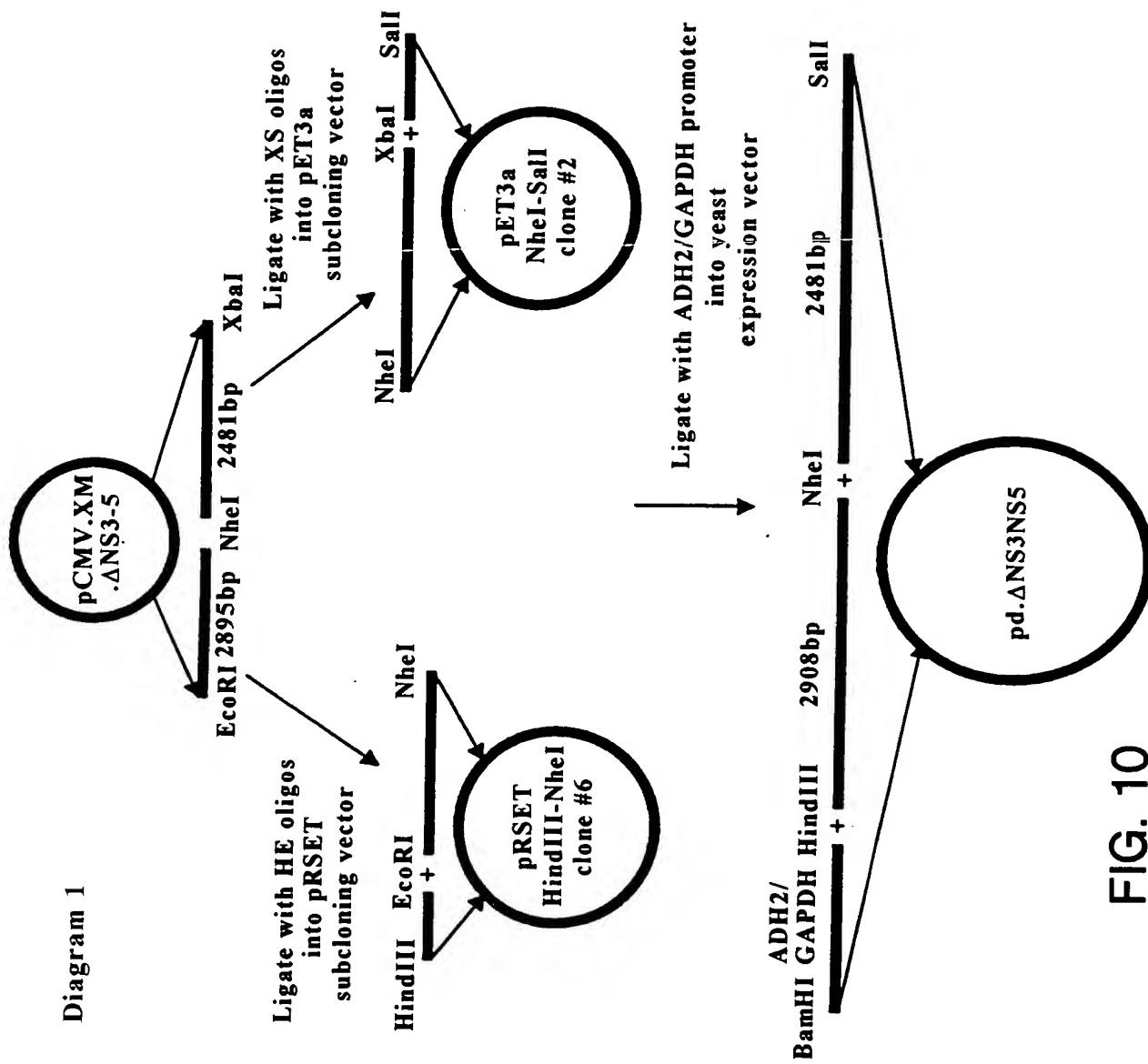


FIG. 10



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal  
2 AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
TCGAATGTTTTGTTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCACGATCAT  
^ ^ ^  
1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,  
  
LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly  
62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG  
GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC  
  
IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr  
122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTAC  
TAGCTAGGATTGTAGTCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG  
^  
122 CLAI,  
  
SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle  
182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA  
AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT  
  
IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu  
242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT  
TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA  
  
AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly  
302 GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC  
CTGGTTCGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCC  
^  
309 ALWN1,  
  
SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle  
362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATC  
AGGCAGTGACACGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG  
  
ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe  
422 CCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC  
GGAAAATGCCGTTCCGATAGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG  
  
CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn  
482 TGTCATTCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTTCGCATTGGGCATCAAT  
ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA  
  
AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal  
542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGCGATGTTGTC  
CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG  
^ ^  
556 SAC2, 566 DRD1,  
  
ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp  
602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC  
CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
^  
621 BSPH1,  
  
CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

FIG. 11-1





662 TGCAATACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAG  
ACGTTATGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTC

722 ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG  
TGTTAGTGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

782 GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
GGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTTCGAC  
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG  
822 BGLI, 839 DRD1,

842 SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
TCGTCCGTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC  
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGG  
887 SACI,

902 GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC  
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTG  
937 SMAI XMAI,

962 HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA  
GTAGAACTTAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT  
991 STUI,

1022 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTG  
AGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCAC  
1075 DRA3,

1082 CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
TGCCTAGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTTCGATTCGC  
ACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTCACAACTAAGCG

1142 LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
CTCAAGCCCACCCCTCCATGGGGCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAGAA  
GAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA  
1156 NCOI,

1202 GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
GAAATCACCTGACGCACCCAGTCACCAATACATCATGACATGCATGTCGGCCGACCTG  
CTTTAGTGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC  
1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

1262 GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
GAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTAT  
CTCCAGCAGTGCTCGTGACCCACGAGCAACCGCCGAGGACCGACGAACCGGCGCATA

FIG. 11-2



1322 CysLeuSerThrGlyCysValValIleValGlyArgValValL uSerGlyLysPr Ala  
TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA  
ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT  
1375 NAEI,

1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAG  
TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC  
1391 DRD1,

1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC  
GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG

1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAAC  
CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
1508 PSTI, 1513 TTH3I,

1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
TGGCAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAA  
ACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTGAAGTAGTCACCCTATGTT  
1571 XHOI, 1592 NDEI,

1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
TACTTGGCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT  
ATGAACCGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACACCGAAAA  
1649 BSTE2,

1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
ACAGCTGCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGG  
TGTCGACGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCC  
1683 ALWN1 PVU2,

1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
GGGTGGGTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC  
CCCACCCACCGACGGGTGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCG  
1800 ESP1,

1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
TTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA  
AATCGACCGCGGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT  
1808 KAS1 NARI,

1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC  
CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG

FIG. 11-3



- 1884 SACI, 1905 BSPH1,
- 1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaL uVal  
CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTA  
GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT  
^
- 1934 TTH3I,
- 1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTG  
CAGCCGCACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCAC  
^ ^
- 2010 NAEI, 2023 SMAI XMAI,
- 2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCACGCAC  
GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG  
^ ^
- 2073 SMAI XMAI, 2099 DRA3,
- 2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA  
ATGCACGGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACAT  
^
- 2121 PVU2,
- 2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
ACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCTCCATGCTCC  
TGGGTGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGG  
^ ^
- 2165 ALWN1, 2170 MST2,
- 2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
GGTTCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC  
CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGG  
^
- 2226 ECON1,
- 2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTGTGTCTGCCAGCGC  
ACCGATTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCG  
^ ^ ^
- 2291 ESP1, 2306 PVU2, 2316 BAMHI,
- 2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA
- 2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGG  
CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC  
^ ^ ^
- 2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,
- 2462 AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyPr CysThrProLeu  
AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCCTGTACCCCCCTT  
TTGTACACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAA

FIG. 11-4



2486 ASE1, 2503 APAI,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGCTTGATGTGCAAGCGGATACCTCCCACAGACGTCTCCTTATGCACCTCTAT  
^

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCC  
TCCGTCCACCCCCTGAAGGTGATGACTGCCCATCTGATGACTGTTAGAATTTACGGGC  
^

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCCCGAATTTTTTCACAGAATTGGACGGGGTGCCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAA  
CGCGGGGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGG  
^

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGA  
TACGAGTGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCCT  
^

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA  
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGT  
^

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGCACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC  
^

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG  
TCCGTCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT  
^

3102 BGL2,

FIG. 11-5



3122 IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr  
ATCCTGCGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTAT  
TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATA  
3149 ALWN1, 3170 EAG1 XMA3,  
3182 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC  
TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG  
3223 HGIE2, 3235 NCOI,  
3242 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal  
TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG  
ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC  
3302 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC  
CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCG  
3338 SACI, 3352 HIND3,  
3362 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro  
AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCT  
TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA  
3422 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
TCTGGCTGCCCCCGACTCCGACGCTGAGTCTTATCCTCCATGCCCCCTGGAGGGG  
AGACCGACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC  
3443 EAM11051,  
3482 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
3490 BAMHI, 3491 BSAB1, 3493 BSPE1,  
3542 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
GCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCG  
CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC  
3595 DRA3,  
3602 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
TGCGCCGCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC  
ACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTAGCAACGATGCAGTG  
3606 SAC2, 3617 ALWN1, 3661 PFLM1,  
3662 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
CACAAATTTGGTGTATTCCACCACCTCACGAGTGCTTGCCAAAGGCAGAAAGTACACA  
GTGTTAAACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGT  
3687 DRA3,  
PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

FIG. 11-6



3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA  
AAACTGTCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGT

3782 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCC  
CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGG

3822 HIND3,

3842 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTGTCATGCCAGA  
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCT

3881 AAT2, 3896 BGLI,

3902 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA  
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

3962 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT  
TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCA

4022 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
CGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG  
GCATTCCGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC

4082 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
GCTTTGTACGACGTGGTTACAAAGCTCCCCCTGGCCGTGATGGGAAGCTCCTACGGATTCT  
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAG

4142 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr  
CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC  
GTTATGAGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGG

4166 ECORI,

4202 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC  
GGTTACCCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGAATCTCGCTGTAG

4235 DRD1, 4242 ALWN1,

4262 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATC  
GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCCGGCGCACCGGTAG

4307 BGLI, 4314 BALI,

4322 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
AAGTCCCTCACCAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC  
TTCAGGGAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCTCTTG

4351 APAI,

4382 CysGlyTyrArgArgCysArgAlaS rGlyValLeuThrThrSerCysGlyAsnThrLeu  
TGCGGCTATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTC

FIG. 11-7



ACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAG

4442 ThrCysTyrI1 LysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet  
ACTTGCTACATCAAGGCCCGGGCAGCCTGTGCGAGCCGCAGGGCTCCAGGACTGCACCATG  
TGAACGATGTAGTTCCGGGCGCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTAC  
^

4458 SMAI XMAI,

4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla  
CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCG  
GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGC  
^ ^

4514 DRD1, 4517 TTH3I,

4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro  
GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCC  
CGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGG

4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla  
CCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCC  
GGTGTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG  
^

4643 SACI,

4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla  
CACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCG  
GTGCTGCCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGC  
^

4737 NRUI,

4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle  
AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATC  
TCTCGACGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAG

4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu  
ATGTTTGGCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTT  
TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAA  
^^

4812 PFLM1, 4813 DRA3,

4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer  
ATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC  
TATCGGTCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG  
^

4899 BGL2,

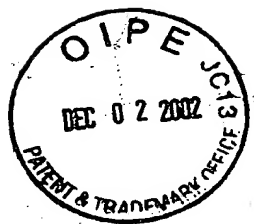
4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer  
ATAGAACCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCA  
TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGT  
^

4960 NCOI,

4982 LeuHisSerTyrSerPr GlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly  
CTCCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGG  
GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCCACGGCGTACGGAGTCTTTTGAACCC  
^

5021 SPHI, 5041 KPNI,

FIG. 11-8



5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
GTACCGCCCTTGGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC  
CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG  
5070 APAI, 5097 BALI,  
ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG  
TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTC  
5119 NDEI,  
LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
5162 CTCAAACTCACTCCAATAGCGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT  
GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA  
5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,  
GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
5222 GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGG  
CCGATGTGCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGCGACCTAGACC  
5246 DRA3,  
PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 TTTTGCCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGG  
AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTACTTCC  
5301 PSTI, 5331 HGIE2,  
5342 TTGGGGTAAACACTCCGGCCTAAAAAATACTAGAACCCGAGTCGAC  
AACCCCATTTGTGAGGCCGATTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG  
5378 XBAI, 5390 SALI,

FIG. 11-9



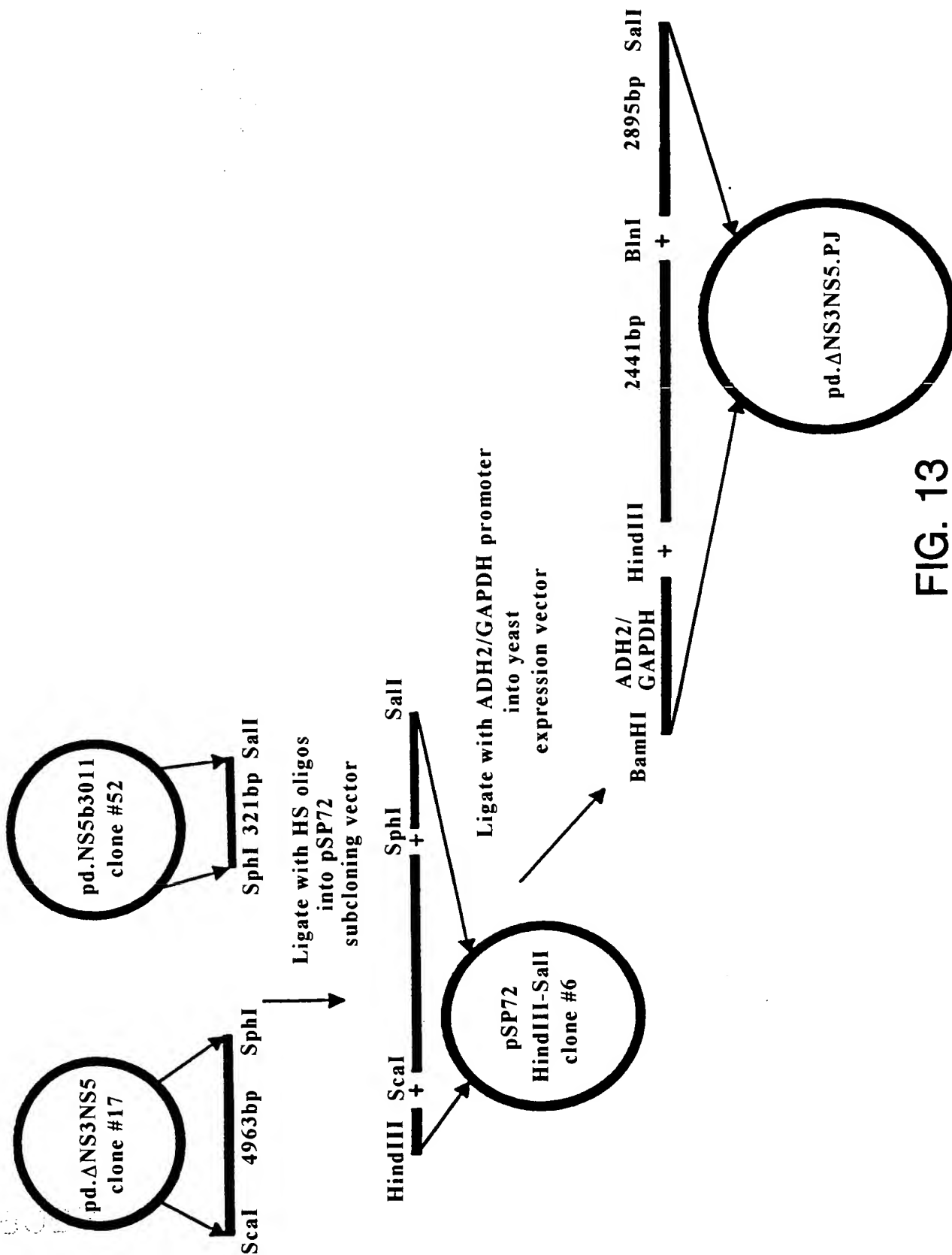


FIG. 13



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACCTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCCCGCTACAACAGCAGCAC  
^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,  
  
ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

FIG. 14-1



662 ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTCTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACCTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

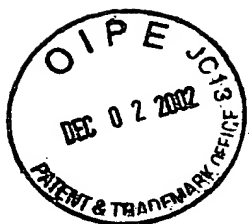
1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTGAATC  
GGGTGGGAGGTACCCGGTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

FIG. 14-2



1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^  
1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^  
1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^  
1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^  
1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA  
^ ^  
1643 BSTE2, 1677 ALWN1 PVU2,

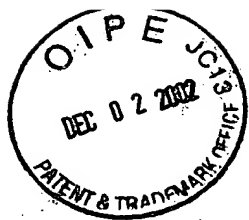
1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^  
1794 ESI1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGACGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
^  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^  
1878 SACI, 1899 BSPH1,

FIG. 14-3



1922 ThrGluAspLeuValAsnLeuLeuPr AlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCAGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCTGTACCCCCCTTCTGCG  
ACCTACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,  
ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

FIG. 14-4



2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACC GGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAAC TGACGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGCGGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGTTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTGTTTCAACCTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

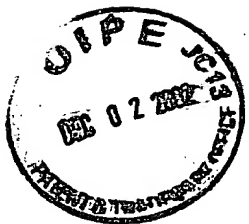
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgPr AspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC

FIG. 14-5



GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACTGTGGTCCATGGCTGCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
3242 CTTCACCTCCAAAGTCCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
3302 ACTGAATCAACCTATCTACTGCCTTGCCGAGCTCGCCACCAGAAGCTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGCTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCCGTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTTGGGGCACGCGG  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTTCAGTGTAAGCTG  
3681 DRA3,  
ArgLeuGlnValLeuAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGTCATGAGTTCTTCCAATTTTCGTCGCCGC

FIG. 14-6



3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGTGTG  
^

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
^ ^

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
^

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
^ ^

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAAATGGTTAAGTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAAGTCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

FIG. 14-7





4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyL uGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCGGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
^  
4452 SMAI XMAI,  
  
4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
^ ^  
4508 DRD1, 4511 TTH3I,  
  
4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT  
  
4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^  
4637 SACI,  
  
4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^  
4731 NRUI,  
  
4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
  
4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCCAGCTGTGGGCGAGGATGATACTGATGACCCATTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^  
4806 PFLM1, 4807 DRA3,  
  
4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACCTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^  
4893 BGL2,  
  
4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^  
4954 NCOI,  
  
4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^ ^  
5015 SPHI, 5035 KPNI,  
  
ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

FIG. 14-8



5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAATAGTCGAC  
GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTATCAGCTG  
5295 PSTI, 5336 SALI,

FIG. 14-9

RECEIVED

DEC 11 1993

FROM GENETIC ENGINEERING

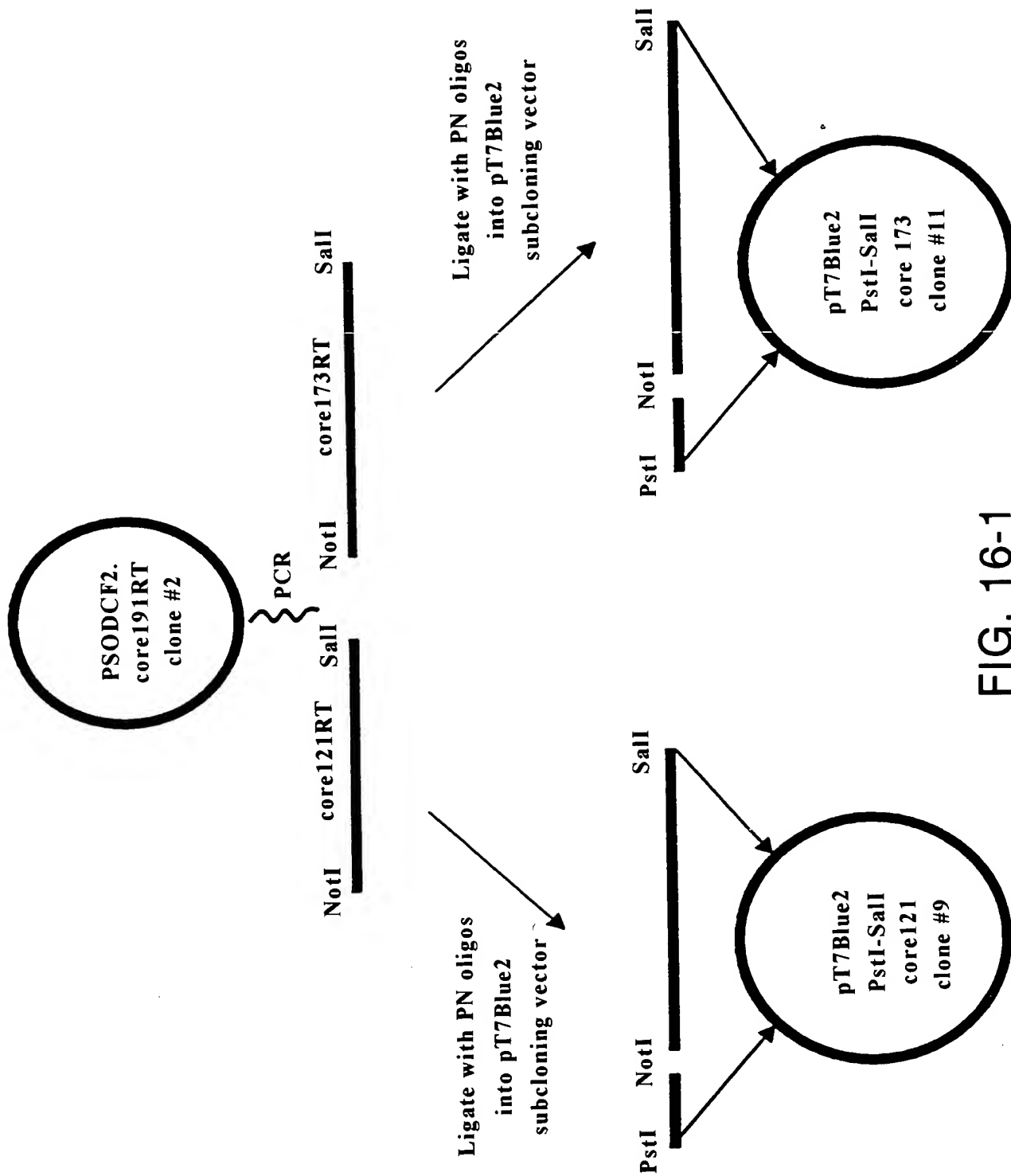


FIG. 16-1

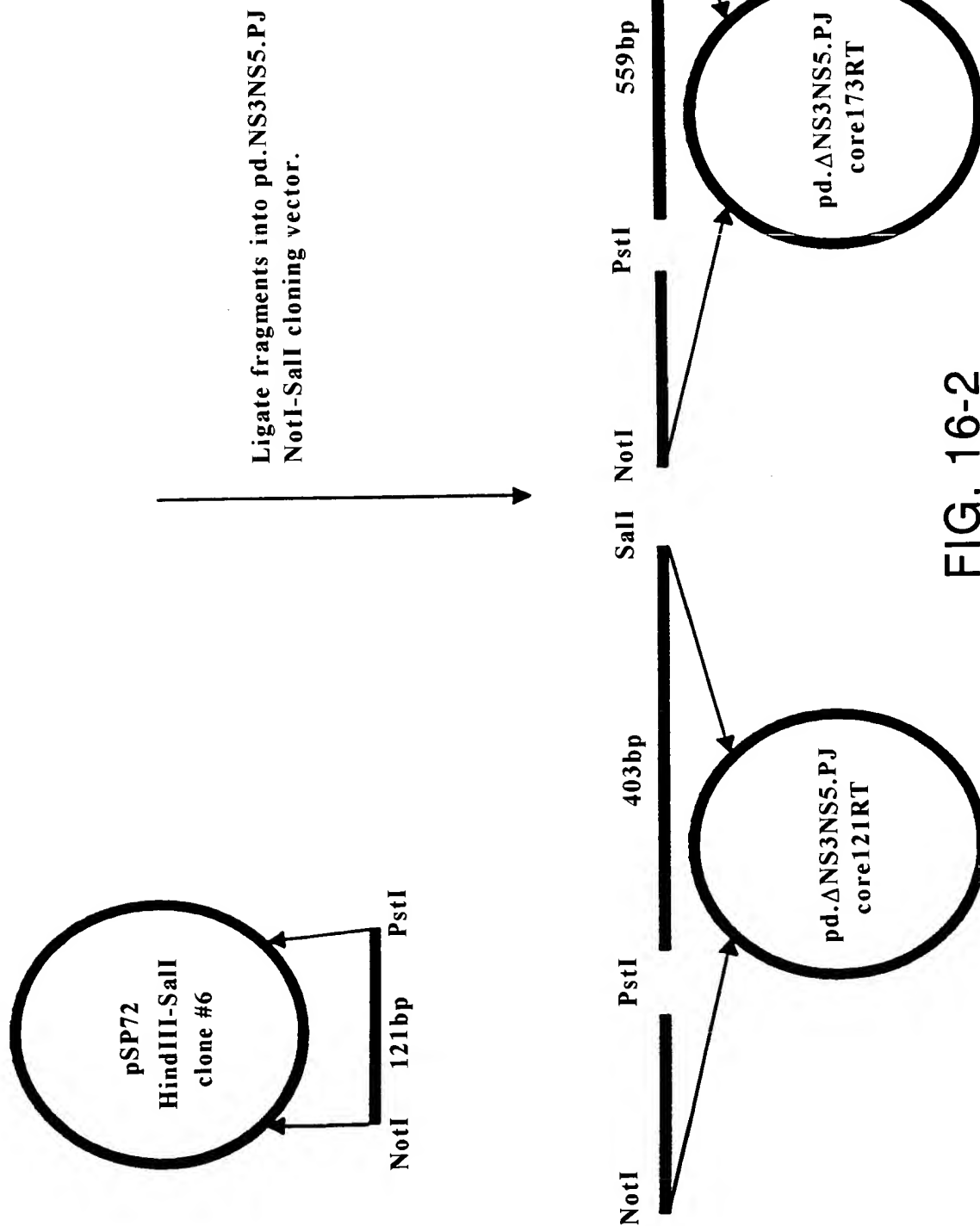


FIG. 16-2



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 17-1

RECEIVED

DEC 18 2003

TECH CENTER 18015000



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIle AsnAlaVal  
TCAAAGAAGAAGTGCAGCACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCAACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

FIG. 17-2



TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTGAGAATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGATCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAAGTACCGAAAATGTGCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTACCAGCCCACTAACCCTAGCCAAACCCTCCTCTCAACATATTGGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 17-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^

1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGACAGGGTAT  
CCGCGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
^

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^

1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
^

1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
^ ^

2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
^ ^

2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGT  
^ ^

2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
^ ^

2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuM tProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCGAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT  
^ ^ ^

2285 ESP1, 2300 PVU2, 2310 BAMHI,







2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^  
3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGCTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCATAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 17-6



3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGT  
3681 DRA3,  
3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGTCGCCG  
3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
3816 HIND3,  
3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,  
3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC  
4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGACGCGCACACGTTTTCTACCGAAAC  
4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG  
4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,  
4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

FIG. 17-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCCGGGCAGCCTGTGCGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTGCGACAGCTCGGCGTCCCGAGGTCTTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCGTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGAGACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIle TyrGlyAlaCysTyrSerIle Glu

FIG. 17-8



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCACGCTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCCGCGCAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGATTGTGGTTGGCCGCCGGCGTCCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuL uProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTTCGCGCGAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 17-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM  
CCTAGCTGGGGCCCCACAGACCCCGCGTAGGTCGCGCAATTTGGGTAAGTAATAGTCG  
GGATCGACCCCGGGGTGTCTGGGGCCGCATCCAGCGCGTTAAACCCATTATTATCAGC  
5650 APAI, 5698 SALI,

5702 AC  
TG

FIG. 17-10

RECEIVED

DEC 11 2002

TECH. CENTER, BOSTON



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCTTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGGCATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,

FIG. 18-1



662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTACCCAGACAGTCGATTTACGCCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGTAAGTCTGTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTTCGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGL1, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCACTACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCAG  
CTTAAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC  
GGGTGGGAGGTACCCGGTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrS rThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG





CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^

1794 ESI1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
^

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaL uValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCCGCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGGAGG  
^ ^

1878 SACI, 1899 BSPH1,

FIG. 18-3



1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTCCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
2342 AAGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTA  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCTGTACCCCCCTTCCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,

FIG. 18-4



2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCCGAATTTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACC GGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACGTCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGCGGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTTACCCGCCGTTGTAGTGTTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC

3096 BGL2,

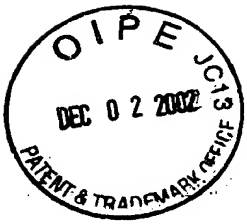
ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

FIG. 18-5



3122 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
3242 CTTCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
3302 ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG  
3681 DRA3,  
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG

FIG. 18-6



TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGTCGCCG

3782 SerLysValLysAlaAsnLeuLeuS rValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCGCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAAGTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

FIG. 18-7



4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
^  
4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
^ ^  
4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
^

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^  
4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^  
4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^  
4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTTCGAACCTTGTCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^  
4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
^  
4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGTACGGAGTCTTTTGAACCCCATGGC  
^ ^  
5015 SPHI, 5035 KPNI,

FIG. 18-8



5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
5113 NDEI,  
5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCGGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
5295 PSTI,  
5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGCGTCTCTGCAGTTCAAG  
5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,  
5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
5449 APAI,  
5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,  
5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,  
5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC



5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA  
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr  
GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTAT  
CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATA  
5772 BSTXI, 5775 APAI,

5822 AlaThrGlyAsnLeuProGlyCysSerOC AM  
GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC  
CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG  
5854 SALI,

FIG. 18-10



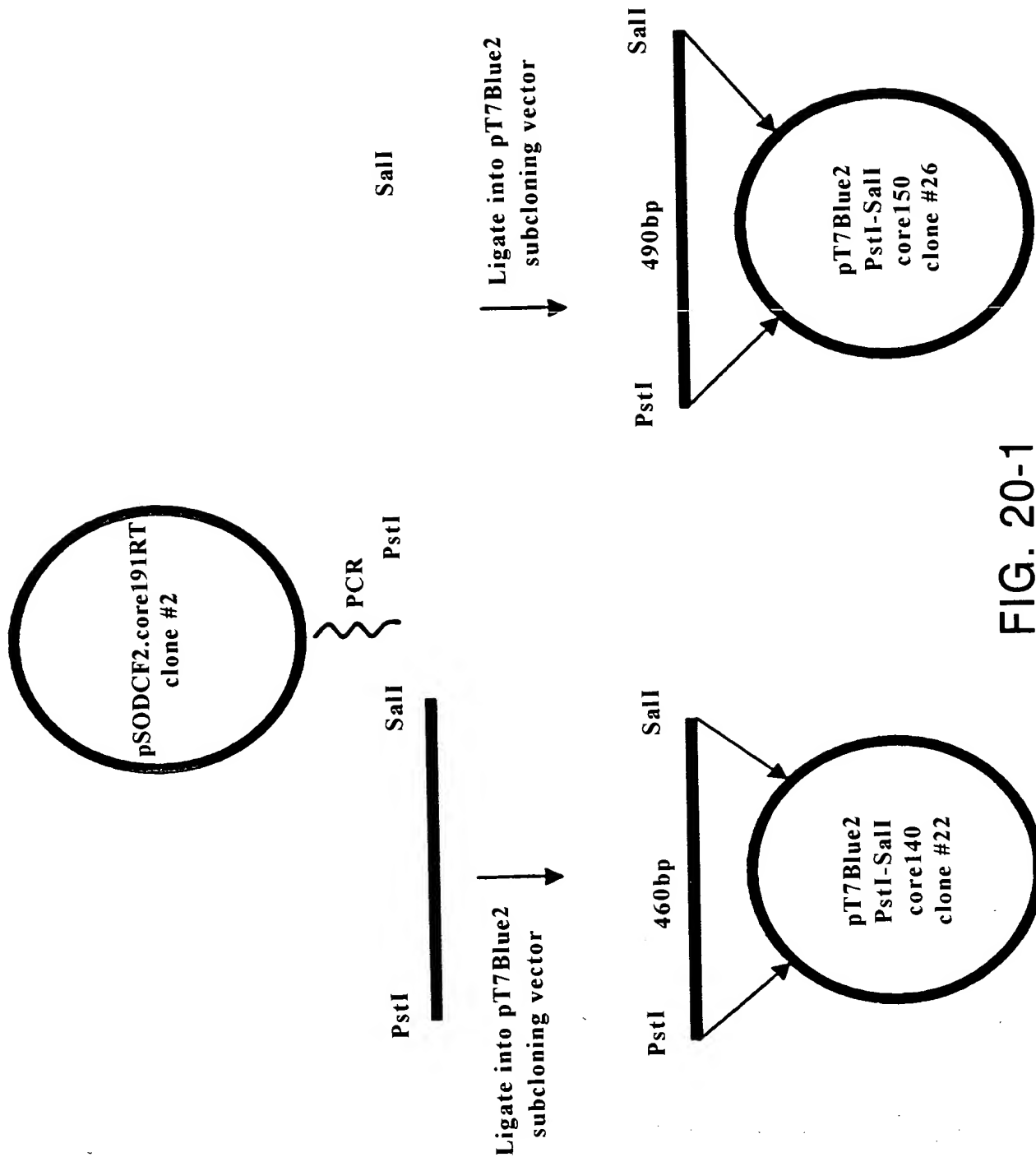


FIG. 20-1

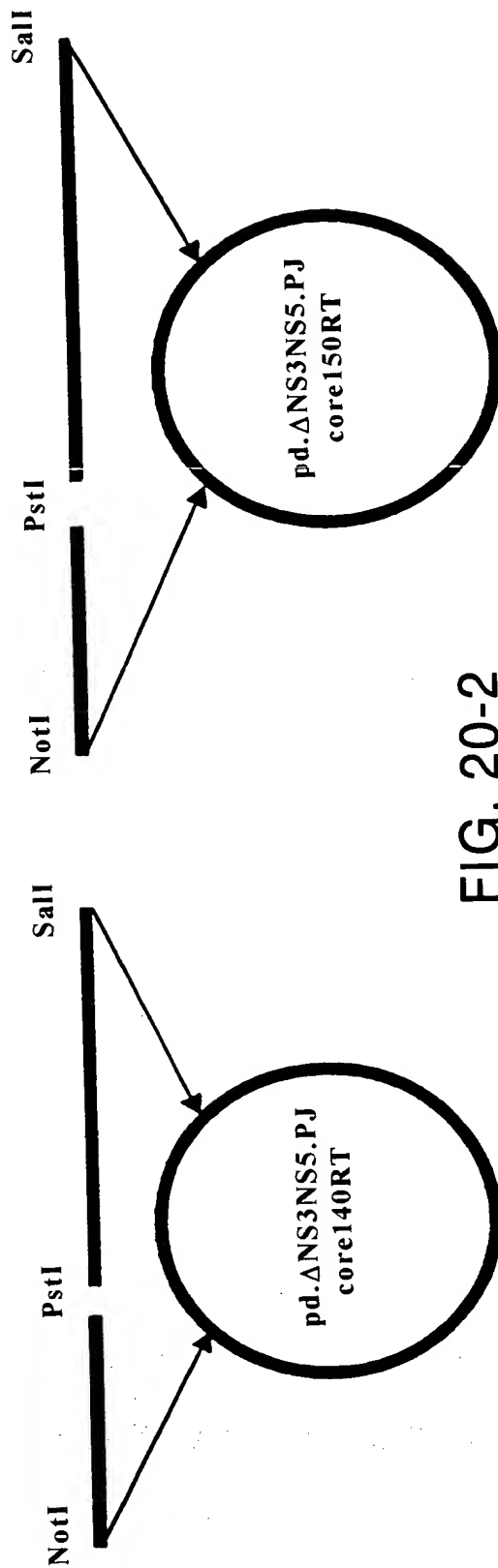
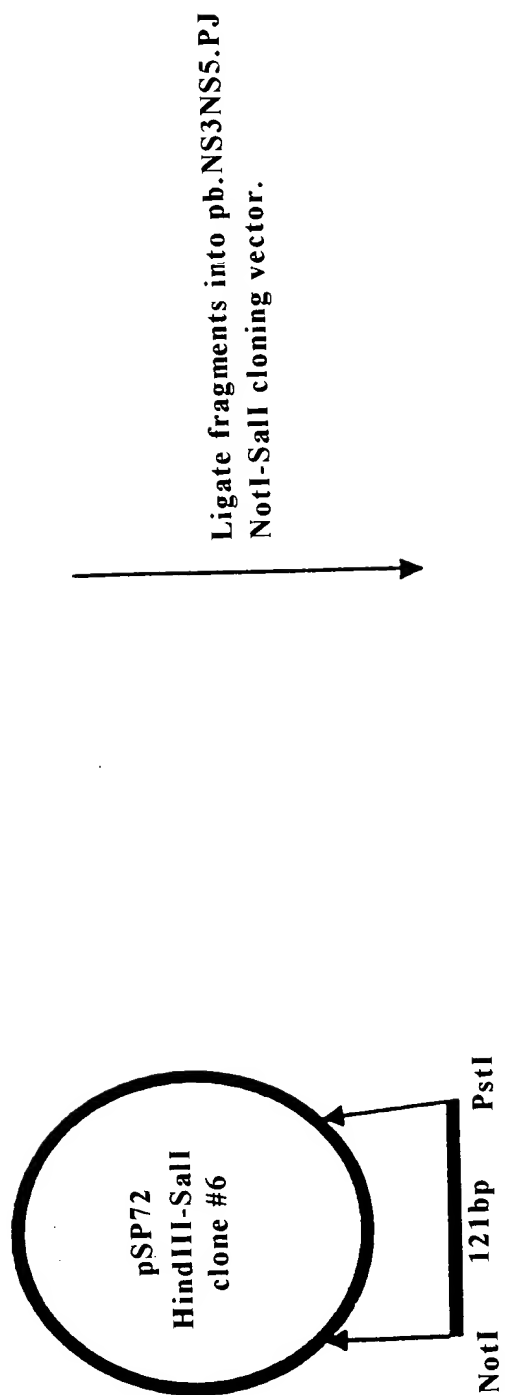
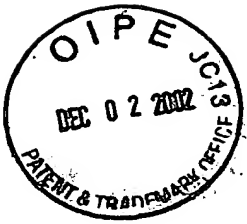


FIG. 20-2



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 21-1

RECEIVED  
DEC 02 2002  
OIGS/DOJ/RECEIVED



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCTCGTCGTG  
CGGATGATGGCGCCAGAAGTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGL1, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SAC1,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGTTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCCTTCTATCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTTCGTCTACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysL uileArgLeuLys  
AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTGCGCTCAAG

FIG. 21-2



TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyPr ThrPr LeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCTCAGAATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGATCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACCTACCGAAAATGTGCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 21-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCGGTGGCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGTAT  
CCGCGGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCATA  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGCGGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCCGACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCGCTCCCCCGTCACGTACCC  
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTCCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTTCGCGCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 21-4



2342 LysGlyValTrpArgGlyAspGlyIleM tHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
^ ^ ^  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC  
^ ^  
2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACCTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^  
2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^  
2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^  
2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG  
^  
2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGCGCCGCTTCCAACCGCTCCCCTAGTGGG  
^  
2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^  
2889 BALI, 2903 NHEI,

FIG. 21-5



2942 ThrAlaAsnHisAspSerPr AspAlaGluLeuIl GluAlaAsnLeuL uTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCAGGCCCTGCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^  
3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrS rTrpThrGlyAlaLeuValThrProCysAla  
GATGTGCTGTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 21-6





3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCCGTAATGGTCCTGCATGAGTTCCTCCAATTTTCGTCGCCCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGTTTTCTACCGAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTGAGTACTCTCGCTGTAGGCATGC

FIG. 21-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCTGAGTCCGCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTGTTCTGTGTGAGGTCTAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 21-8



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCTGAACCTTGTCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAGCTCAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCAATTGTGGTTGGCCGCCGGCGTCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 21-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGGTTAAACCCATTCCAGTAGCTA  
^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG  
^ ^

5724 HGIE2, 5755 SALI,

FIG. 21-10



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 22-1



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAAGTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGTTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCGCCTCAAG

FIG. 22-2



TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluI  
CCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCTGAGTAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGTACTCTACCTTCTCAGAGAGTCGTGAAT

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCTCCTCTTCAACATATTGGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 22-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPh ValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^  
1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
^  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^  
1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
^  
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGTGAACCGGGCCGCTCCCCCGTCACGTCACC  
^ ^  
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCAGCCTACGTCG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC  
^ ^  
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTC  
^ ^  
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
^ ^  
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCTCTGCCAGCGGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCCATA  
^ ^ ^  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 22-4





2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
^ ^ ^

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG  
ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC  
^

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACCTACAGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCAGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACATGCAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGCGCCGCTTCCAACCGCTCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerS rAlaSerGlnLeuSerAlaProSerL uLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^

2889 BALI, 2903 NHEI,

FIG. 22-5



2942 ThrAlaAsnHisAspSerPr AspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,  
3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,  
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCAGGCCCTGCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,  
3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGG  
^ ^  
3217 HGIE2, 3229 NCOI,  
3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,  
3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,  
3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
3542 AspValValCysCysS rMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 22-6



3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGT

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATT

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTGCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTGGGGTTAC

4160 ECORI,

4202 GlyPh SerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCAAGAGCATACTATGGGCGACGAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC

FIG. 22-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCTCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCCGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGCGGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

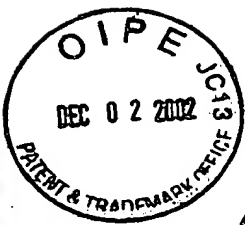
4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 22-8



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^ ^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^ ^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGCACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGTTTGCATTGTGGTTGGCCGCCGGCGTCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 22-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 KHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM  
GGAGGCGCTGCCAGGGCCTAATAGTCGAC  
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

FIG. 22-10